

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-117937

(43)Date of publication of application : 27.04.2001

(51)Int.Cl.

G06F 17/30

(21)Application number : 11-297604

(71)Applicant : HITACHI LTD

(22)Date of filing : 20.10.1999

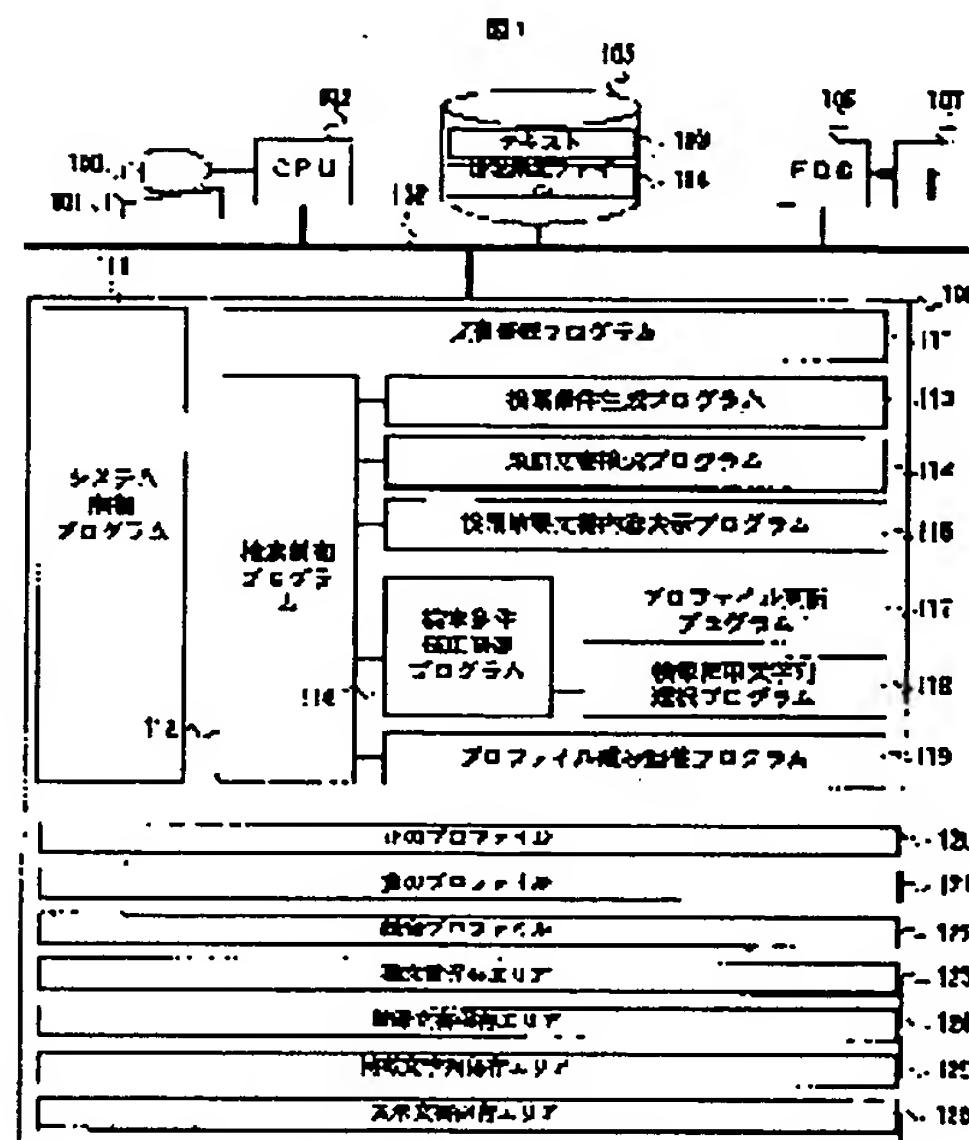
(72)Inventor : INABA YASUHIKO
TADA KATSUMI
SUGAYA NATSUKO
MATSUBAYASHI TADATAKA
YAMAGUCHI AKIHIKO
KAWASHITA YASUSHI

(54) METHOD AND DEVICE FOR RETRIEVING DOCUMENT

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a system capable of easily improving the accuracy of retrieval on the basis of a suitable or unsuitable user evaluation to the retrieved result in the case of similar document retrieval for retrieving the document of contents similar to a document designated by a user.

SOLUTION: Retrieval condition data are updated while using a character string extracted from an evaluation object document, and retrieval is performed while using a character string, which is not contained in one part or all character strings extracted from a document evaluated as desired one, between a character string extracted from a document evaluated desired for the user and a character string extracted from a document evaluated non-desired for the user.



LEGAL STATUS

[Date of request for examination]

19.03.2003

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the
examiner's decision of rejection or application converted
registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of
rejection]

[Date of requesting appeal against examiner's decision of
rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開2001-117937

(P2001-117937A)

(43)公開日 平成13年4月27日 (2001. 4. 27)

(51)Int.Cl.⁷

識別記号

F I

テーマコード(参考)

G 0 6 F 17/30

G 0 6 F 15/403

3 4 0 B

5 B 0 7 5

15/40

3 7 0 A

15/403

3 5 0 C

審査請求 未請求 請求項の数6 O L (全 25 頁)

(21)出願番号

特願平11-297604

(22)出願日

平成11年10月20日 (1999. 10. 20)

(71)出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72)発明者 稲場 靖彦

神奈川県川崎市幸区鹿島田890番地 株式

会社日立製作所システム開発本部内

(72)発明者 多田 勝己

神奈川県川崎市幸区鹿島田890番地 株式

会社日立製作所システム開発本部内

(74)代理人 100075096

弁理士 作田 康夫

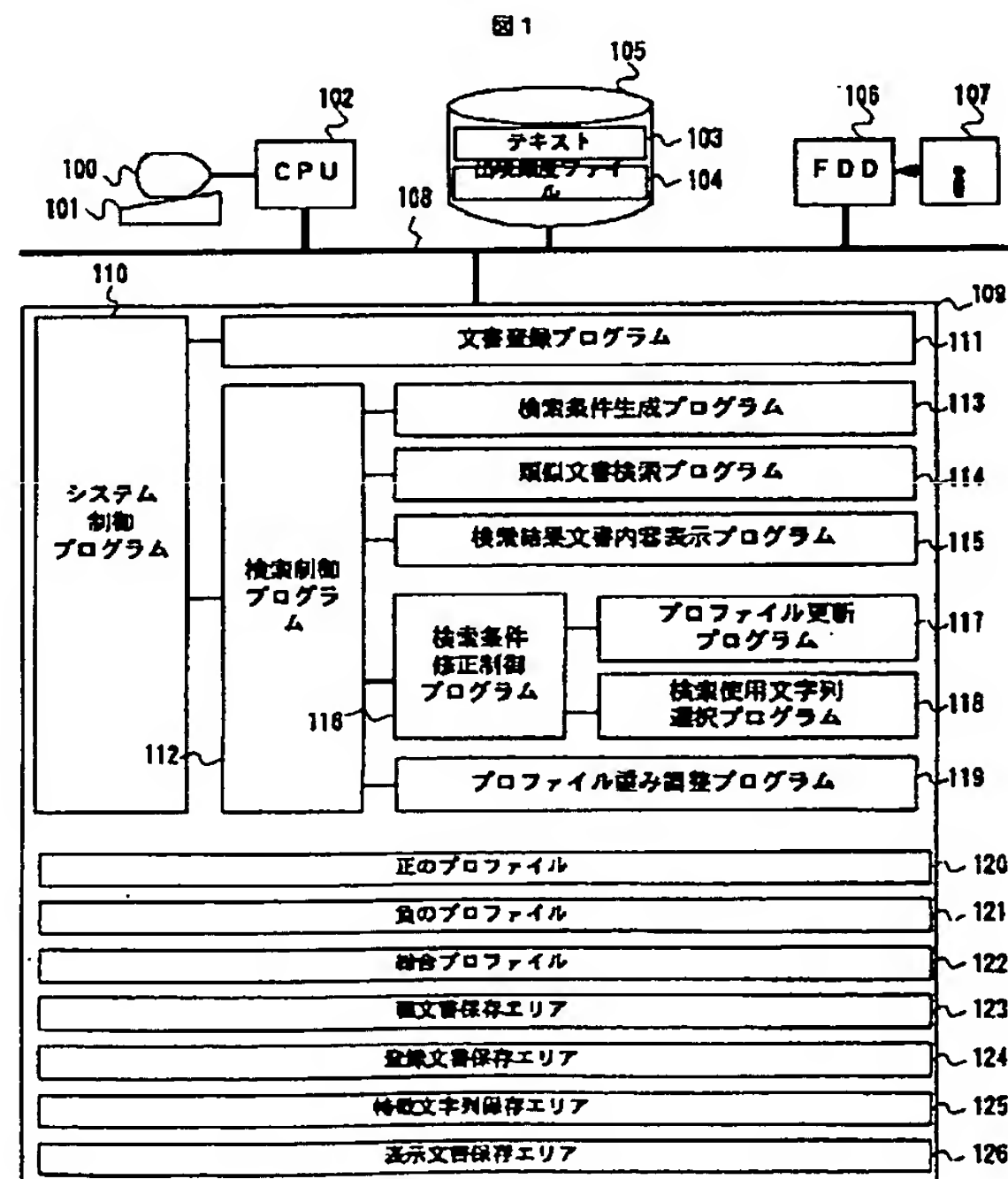
最終頁に続く

(54)【発明の名称】 文書検索方法および装置

(57)【要約】

【課題】ユーザが指定した文書と内容的に類似した文書を検索する類似文書検索において、検索結果に対する適、不適といったユーザ評価にもとづいて、検索の精度を簡易に向上させることのできるシステムを提供する。

【解決手段】評価対象文書から抽出された文字列を用いて検索条件データを更新し、ユーザが所望のものであると評価をした文書から抽出した文字列と、ユーザが所望のものでないと評価した文書から抽出した文字列のうち所望のものであると評価をした文書から抽出した文字列の一部または全部に含まれないもの、を用いて検索を行なう。



【特許請求の範囲】

【請求項1】文字列に付与された重みを含む検索条件により文書データベースを検索し、該検索により得られた文書に対してユーザが入力した「所望である」または「所望でない」の評価を受け取り、上記検索の結果得られた文書から抽出した文字列の重みを上記評価に基づき変更して検索する文書検索方法において、
上記「所望である」と評価した文書から抽出した第一の文字列に正の重みを付与し、
上記「所望でない」と評価した文書から抽出した第二の文字列に負の重みを付与し、
上記第二の文字列のうち上記第一の文字列と一致するとともに当該第一の文字列の重みが所定値以上となるものを除外したものとその重みおよび上記第一の文字列とその重みとを含む検索条件を生成して検索することを特徴とした文書検索方法。

【請求項2】文字列に付与された重みを含む検索条件により文書データベースを検索し、該検索により得られた文書に対してユーザが入力した「所望である」または「所望でない」の評価を受け取り、上記検索の結果得られた文書から抽出した文字列の重みを上記評価に基づき変更して検索する文書検索方法において、
上記「所望である」と評価した文書から第一の文字列を抽出し、
上記「所望でない」と評価した文書から抽出した文字列で上記第一の文字列と一致する場合は、当該第一の文字列の重みが所定値以下の場合は上記抽出した文字列を第二の文字列として抽出し、第二の文字列の重みを第一の文字列の重みよりも低くし、一致しない場合は上記抽出した文字列を第二の文字列として抽出し、第二の文字列の重みを第一の文字列の重みよりも低くすることを特徴とした文書検索方法。

【請求項3】文字列に付与された重みを含む検索条件により文書データベースを検索し、該検索により得られた文書に対してユーザが入力した「所望である」または「所望でない」の評価を受け取り、上記検索の結果得られた文書から抽出した文字列の重みを上記評価に基づき変更して検索する文書検索方法において、
上記「所望である」と評価した文書から第一の文字列を抽出し、
上記「所望でない」と評価した文書から抽出した文字列で上記第一の文字列と一致しない場合は上記抽出した文字列を第二の文字列として抽出し、第二の文字列の重みを第一の文字列の重みよりも低くすることを特徴とした文書検索方法。

【請求項4】請求項1または請求項2記載の文書検索方法において、
上記検索結果文書に対するユーザの評価は、ひとつ以上の段階評価を設定する方法であり、
文字列の重みの変更方法は、前記評価に応じて多段階に

変更の程度を設定する方法であることを特徴とする文書検索方法。

【請求項5】請求項1または請求項2記載の文書検索方法において、ユーザの評価に基づき検索条件を更新する際に、

ユーザが評価した文書から抽出した文字列について、検索条件に用いるか否かをユーザが選択情報を入力して選択することすることを特徴とした文書検索方法。

【請求項6】文字列に付与された重みを含む検索条件により文書データベースを検索し、該検索により得られた文書に対してユーザが入力した「所望である」または「所望でない」の評価を受け取り、上記検索の結果得られた文書から抽出した文字列の重みを上記評価に基づき変更して検索する文書検索装置において、
上記「所望である」と評価した文書から抽出した第一の文字列に正の重みを付与し、

上記「所望でない」と評価した文書から抽出した文字列が上記第一の文字列と一致する場合は当該第一の文字列の重みが所定値以下の場合は上記抽出した文字列を第二の文字列として負の重みを付与し、上記第一の文字列と一致しない場合は上記抽出した文字列を第二の文字列として負の重みを付与する手段と、

上記第一の文字列とその重みおよび上記第二の文字列とその重みとを含む検索条件を生成して検索する手段とを備えたことを特徴とした文書検索方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、検索条件に基づいて文書データベースから文書を検索する方法および装置に関し、その検索の結果として得られた文書に対してユーザが評価を与え、その評価に基づき検索条件を変更する方法および装置に関する。

【0002】

【従来の技術】近年、パーソナルコンピュータやインターネット等の普及に伴い、電子化文書が急激に増加している。このような状況において、ユーザが所望する情報を含んだ文書を高速かつ効率的に検索したいという要求が高まってきている。

【0003】このような要求に応えるための検索技術としてレリバンスフィードバックとよばれる技術がある。この技術は、全文検索や類似文書検索による検索結果に対して、ユーザが「所望の文書である」か「所望の文書でない」かなどの評価をシステムに入力し、その評価情報を検索条件に反映させることにより、その後の検索結果を改善する技術である。

【0004】具体的な処理の内容としては、例えば「"Information Retrieval", William B. Frakes / Rocardo Baeza-Yates, Prentice Hall PTR, 1992 p.p.241~263」に示されるように、ユーザが所望であると評価した文書から抽出した単語に関する検索条件中の重みを加算し、

所望でないと評価した文書から抽出された単語に関する検索条件中の重みを減算する方法がある。以下この技術を従来技術 1 と呼ぶ。検索条件中のある単語について、具体的な重みの加減算の方法の例を式 1 に示す。

【0005】

【数 1】

$$W' = W + \alpha \sum_i^P FP(i) - \beta \sum_j^N FN(j) \quad \dots (数 1)$$

【0006】ここで W' はその単語の新たな重み、 W は元の重みであり、 $FP(i)$ は所望であると評価された i 番目の文書におけるその単語の出現回数、 $FN(j)$ は所望でないと評価された j 番目の文書におけるその単語の出現回数である。また、 P は所望であると評価された文書の数、 N は所望でないと評価された文書の数である。なお、 α 、 β はパラメータである。ここで、この新たな重み W' は負になってもよく、そのような場合は、その単語が含まれる文書は類似度が下がることになる。

【0007】この従来技術 1 によるレリバンスフィードバック処理の例を図 2 に示す。本図に示す例は、ユーザが「高校野球」に関する文書を所望する場合に、「サッカーに続き高校野球が開幕した」という文書を種文書に選んだ場合である。その後、「サッカー」に関するノイズ文書に対し「所望でない」と評価をして、システムに入力した場合である。この結果、本図に示すように「サッカー」という単語の重みが下がり、以後「サッカー」に関する文書の類似度を下げることができる。

【0008】

【発明が解決しようとする課題】しかし、従来技術 1 による方式では、ユーザが「所望のものでない」といった評価をしたときに検索結果が改善しない場合がある。この問題を図 3 を用いて説明する。本図に示した例は、

「高校野球」に関する文書を所望する場合に、「高校サッカーが開幕した・・・」といったノイズ文書に対し「所望の文書でない」と評価した場合である。このとき従来技術 1 によれば、このノイズ文書から「高校」「サッカー」「開幕」といった単語を抽出し、検索条件中のそれぞれの単語の重みを減算することになる。この場合、「サッカー」の重みを減算するだけでなく、「高校」という単語の重みまでも減算してしまう。その結果、更新された検索条件によって検索を行なうと、「高校野球」に関する文書の類似度が、「プロ野球」「社会人野球」といった文書の類似度よりも低くなってしまいう問題がある。

【0009】このように、従来の方法によりユーザが「所望のものでない」と評価した文書から抽出した単語の重みを単純に減算すると、ユーザが所望とする概念を表す単語の重みまで減算してしまい、検索結果が改善しないという問題がある。

【0010】本発明の目的は、ユーザが「所望のものでない」といった評価を与えた文書から抽出した情報のう

ち適切なものを使用して、検索結果を改善することにある。

【0011】

【課題を解決するための手段】上記課題を解決するため、第 1 の手段として、文字列に付与された重みを含む検索条件により文書データベースを検索し、該検索により得られた文書に対してユーザが入力した「所望である」または「所望でない」の評価を受け取り、上記検索の結果得られた文書から抽出した文字列の重みを上記評価に基づき変更して検索する文書検索方法において、上記「所望である」と評価した文書から抽出した第一の文字列に正の重みを付与し、上記「所望でない」と評価した文書から抽出した第二の文字列に負の重みを付与し、第二の文字列のうち上記第一の文字列と一致するもとともに当該第一の文字列の重みが所定値以上ものを除外したものとその重みおよび上記第一の文字列とその重みとを含む検索条件を生成して検索する。

【0012】この方法により、ユーザが所望のものと評価した文書から抽出した所望の内容を特徴付ける文字列に付与された負の重みにより検索精度を下げてしまうという課題を改善することができる。

【0013】また、第 2 の手段は、文字列に付与された重みを含む検索条件により文書データベースを検索し、該検索により得られた文書に対してユーザが入力した「所望である」または「所望でない」の評価を受け取り、上記検索の結果得られた文書から抽出した文字列の重みを上記評価に基づき変更して検索する文書検索方法において、上記「所望である」と評価した文書から第一の文字列を抽出し、上記「所望でない」と評価した文書から抽出した文字列で上記第一の文字列と一致する場合は、当該第一の文字列の重みが所定値以下の場合は上記抽出した文字列を第二の文字列として抽出し、第二の文字列の重みを第一の文字列の重みよりも低くし、一致しない場合は上記抽出した文字列を第二の文字列として抽出し、第二の文字列の重みを第一の文字列の重みよりも低くする。

【0014】この方法により、ユーザが所望のものと評価した文書から抽出した所望の内容を特徴付ける文字列に、負の重みを付与してしまい以降の検索精度を下げてしまうという課題を改善できる。

【0015】

【発明の実施の形態】以下、本発明の第一の実施例について説明する。

【0016】まず、本発明の第一の実施例のシステム構成を図 1 に示す。本実施例におけるシステムは、ディスプレイ 100、キーボード 101、中央演算処理装置 (CPU) 102、磁気ディスク装置 105、フロッピーディスクドライブ (FDD) 106、主メモリ 109 およびこれらを結ぶバス 108 から構成される。

【0017】磁気ディスク装置 105 は二次記憶装置の

一つであり、テキスト103、出現頻度ファイル104が格納される。FDD106を介してフロッピディスク107に格納されている情報が、主メモリ109あるいは磁気ディスク装置105へ読み込まれる。

【0018】主メモリ109には、システム制御プログラム110、文書登録プログラム111、検索制御プログラム112が格納される。検索制御プログラム112は、検索条件生成プログラム113、類似文書検索プログラム114、検索結果文書内容表示プログラム115、検索条件修正制御プログラム116、およびプロファイル重み調整プログラム119で構成される。ここで、検索条件修正制御プログラム116は、プロファイル更新プログラム117、および検索使用文字列選択プログラム118で構成される。

【0019】また、正のプロファイル120、負のプロファイル121、総合プロファイル122、種文書保存エリア123、登録文書保存エリア124、特徴文字列保存エリア125、および表示用文書保存エリア126が同じく主メモリ109に確保される。

【0020】ここで、正のプロファイル120、負のプロファイル121、総合プロファイル122とは後述する図15に示すように、いずれも幾つかの検索文字列とその重みを保持したデータである。正のプロファイル120には、ユーザが所望であると評価した文書から抽出した文字列が格納される。負のプロファイル121には、ユーザが所望のものでないと評価した文書から抽出した文字列が格納される。総合プロファイル122は、正負のプロファイルから選択された検索に用いる文字列が格納される。

【0021】以下に、第一の実施例における、各プログラムの処理手順について説明する。

【0022】まず、システム制御プログラム110の処理手順について図4のPAD (Problem Analysis Diagram) 図を用いて説明する。

【0023】システム制御プログラム110は、まずステップ401においてユーザがキーボードから入力したコマンドを解析する。

【0024】次にステップ402において、このコマンドが文書登録のコマンドであると解析された場合には、ステップ404で文書登録プログラム111を起動して文書の登録を行なう。

【0025】またステップ403において、検索実行のコマンドであると解析された場合には、ステップ405で検索制御プログラム112を起動して文書の検索を行なう。

【0026】以上が、システム制御プログラム110の処理手順である。

【0027】次に、図4に示したステップ404でシステム制御プログラムにより起動される、文書登録プログラム111について図5のPAD図を用いて説明する。

【0028】文書登録プログラム111は、まずステップ501においてD106に挿入されたフロッピディスク107から登録すべき文書データを読み込み、これをテキスト103として磁気ディスク装置105に格納する。文書データは、フロッピディスク107を用いて入力するだけに限らず、通信回線やCD-ROM装置(図1には示していない)等を用いて他の装置から入力するような構成を取ることも可能である。

【0029】次にステップ502で、検索対象文書から抽出される自立語の可能性のある文字列(以下、特徴文字列と呼ぶ)がどの文書に何回出現したかを高速に抽出するためのデータとして、出現頻度ファイル104を各登録対象文書について生成する。ここで出現頻度ファイルの生成方法としては「特開平11-143902号広報」に開示されている出現頻度ファイルの生成方法と同一の方法でも良いし、形態素解析等を用いて各文書中の単語を抽出する方法やニューラルネットワークの学習データを用いた方法でもかまわない。また、単純n-gramを抽出する方法であってもかまわない。

【0030】以上が、文書登録プログラム111の処理手順である。次に、図4に示したステップ405でシステム制御プログラムにより起動される、検索制御プログラム112の処理手順を図6のPAD図を用いて説明する。

【0031】検索制御プログラム112は、まずステップ601において検索条件生成プログラム113を起動し、検索条件を生成する。

【0032】次にステップ602において、ステップ603～ステップ612の処理を、ステップ604においてユーザから検索セッションの終了が要求されたと解析されるまで繰り返す。

【0033】この繰り返し処理では、まずステップ603において、類似文書検索プログラム114を起動し、ステップ601で生成された検索条件にもとづき類似文書検索を行なう。

【0034】次にステップ604において、キーボードから入力されるコマンドを解析する。

【0035】次にステップ605において、このコマンドが文書の内容表示コマンドであると解析された場合には、ステップ609で検索結果文書内容表示プログラム115を起動し、指定された検索結果文書の内容を表示する。

【0036】次にステップ606において、検索結果文書に対するユーザの評価の入力コマンドであると解析された場合には、ステップ610で検索条件修正制御プログラム116を起動し、検索条件を修正する。

【0037】次にステップ607において、プロファイルの内容調整コマンドであると解析された場合には、ステップ611でプロファイル重み調整プログラム119を起動し、プロファイルの内容を調整する。

【0038】次にステップ608において、検索セッション終了コマンドであると解析された場合には、ステップ612で、正のプロファイル120、負のプロファイル121、および総合プロファイル122の内容をクリアし、ステップ602の繰り返しを終了する。

【0039】以上が検索制御プログラム112の処理手順である。

【0040】次に、図6に示したステップ601で検索制御プログラムにより起動される、検索条件生成プログラム113の処理手順を図7のPAD図を用いて説明する。

【0041】検索条件生成プログラム113は、まずステップ701において、キーボード101から入力される種文書を読み込み、種文書保存エリア123に格納する。

【0042】次にステップ702において、種文書保存エリア123に格納された種文書から特徴文字列を抽出し、種文書内出現回数を計数して、特徴文字列保存エリア125に格納する。

【0043】ここで、特徴文字列を抽出する方法は、図5に示した文書登録プログラム111のステップ502における方法を用いても良いし、その他の方法を用いても良い。

【0044】次にステップ703において、ステップ702で抽出した特徴文字列をステップ702で計数した出現回数と共に総合プロファイル122に書き込む。ここで総合プロファイル122は、後述する図15に示すように特徴文字列とその重みが保持されたものであり、後述するように類似文書検索プログラム114の入力として使用する。ここで重みとしては種文書内出現回数を用いるものとするが、他のものを用いても良い。また、ここで総合プロファイル122に書き込む文字列は、ステップ702で抽出した特徴文字列のうち重みの上位から所定数のものに限定しても良い。

【0045】次にステップ704において、ステップ702で抽出した文字列をステップ702で計数した出現回数と共に正のプロファイル120に書き込む。この正のプロファイル120は、後述するように、検索結果文書に対しユーザが評価をした場合に、検索条件を修正する際に使用する。また、ここで正のプロファイル120に書き込む文字列は、ステップ702で抽出した特徴文字列のうち重みの上位のもの所定数に限定しても良い。

【0046】以上が、検索条件生成プログラム113の処理手順である。

【0047】次に、図6に示したステップ603で検索制御プログラムにより起動される、類似文書検索プログラム114の処理手順を図8のPAD図を用いて説明する。

【0048】類似文書検索プログラム114は、まずステップ801において、図7に示したステップ703で

検索条件生成プログラム113により生成された総合プロファイル122を読み込む。

【0049】次にステップ802において、出現頻度ファイル104を読み込む。

【0050】次にステップ803において、総合プロファイル122内の特徴文字列の重みと、出現頻度ファイル104内の各文書における該文字列の出現頻度から、テキスト103内の各文書の類似度を算出する。ここで類似度の算出式としては、例えば以下の式2のようなものを用いる。

【0051】

【数2】

$$S(D) = \sum_i^N \{ \text{Frq}(i) \times w(i) \} \quad \cdots \text{(数2)}$$

【0052】この式で、S(D)はテキスト103内の文書番号Dの類似度であり、Frq(i)は出現頻度ファイル104内の単語iの文書Dにおける出現頻度であり、w(i)は総合プロファイル内の単語iの重みである。ここで類似度算出式としては、これ以外のものを用いても構わない。

【0053】次にステップ804において、テキスト103内の各文書の文書番号を類似度の順に降順にソートし、ディスプレイ100に出力する。ここで、類似度の上位所定数のみを出力するようにしても良いし、所定の類似度を上回るもののみを出力するようにしても良い。また、文書にタイトルのような属性があればそれを出力しても良い。

【0054】以上が、類似文書検索プログラム114の処理手順である。

【0055】次に、図6に示したステップ609で検索制御プログラムにより起動される、検索結果文書内容表示プログラム115の処理手順を図9のPAD図を用いて説明する。

【0056】検索結果文書内容表示プログラム115は、まずステップ901において、ユーザがキーボード101から入力する文書番号を読み込む。

【0057】次にステップ902において、ステップ901で入力された文書番号に該当する文書を登録文書保存エリア124に読み込む。

【0058】次にステップ903において、ステップ904で該文書を最後まで読み込むまで以下に示すステップ904からステップ907の処理を繰り返す。

【0059】ステップ903の繰り返し処理では、まずステップ904において、登録文書保存エリア124の文書の文字列を順次読み込み、総合プロファイル122に格納された文字列と照合する。

【0060】次にステップ905において、ステップ904で読み込んだ文字列が総合プロファイル122において正の重みを持つ文字列と一致した場合には、ステップ908で「該文字列を赤色表示する」という情報を付

与して表示用文書保存エリア 126 に追加する。ここで例えば HTML (HyperText Markup Language) の形式で表示する場合は、該文字列の前後に赤色表示を表すタグを挿入し、表示用文書保存エリア 126 に追加する。ここで、重みが所定値以下の文字列や、重みの上位所定件に含まれないものは、この処理の対象外にするなどしても構わない。また、表示色は別の色を用いても構わない。

【0061】次にステップ 906 において、ステップ 904 で読み込んだ文字列が総合プロファイル 122 において負の重みを持つ文字列と一致した場合には、ステップ 909 で「該文字列を青色表示する」という情報を付与して表示用文書保存エリア 126 に追加する。ここで例えば HTML の形式で表示する場合は、該文字列の前後に青色表示を表すタグを挿入し、表示用文書保存エリア 126 に追加する。ここで、重みが所定値以下の文字列や、重みの上位所定件に含まれないものは、この処理の対象外にするなどしても構わない。また、表示色はステップ 908 で指定する色以外の別の色を用いても構わない。

【0062】次にステップ 907 において、ステップ 904 で読み込んだ文字列が総合プロファイル内の文字列と一致しない場合には、ステップ 910 で「該文字列を黒色表示する」という情報を付与して表示用文書保存エリア 126 に追加する。ここで例えば HTML の形式で表示する場合は、該文字列の前後に黒色表示を表すタグを挿入し、表示用文書保存エリア 126 に追加する。ここで、表示色はステップ 908、909 で指定する以外の別の色を用いても構わない。

【0063】次にステップ 911 において、表示用文書保存エリア 126 に保存された内容をディスプレイ 100 に表示する。

【0064】以上が、検索結果文書内容表示プログラム 115 の処理手順である。

【0065】次に、図 6 に示したステップ 610 で検索制御プログラムにより起動される、検索条件修正制御プログラム 116 の処理手順を図 10 の PAD 図を用いて説明する。

【0066】検索条件修正制御プログラム 116 は、まずステップ 1001 においてプロファイル更新プログラム 117 を起動し、正のプロファイル 120 および負のプロファイル 121 の内容を更新する。

【0067】次にステップ 1002 において、検索使用文字列選択プログラム 118 を起動し、ステップ 1001 で更新された正のプロファイル 120 および負のプロファイル 121 の内容にもとづき、総合プロファイル 122 の内容を更新する。

【0068】以上が検索条件修正プログラム 116 の処理手順である。

【0069】次に、図 6 に示したステップ 611 で検索

制御プログラムにより起動される、プロファイル重み調整プログラム 119 の処理手順を図 11 の PAD 図を用いて説明する。

【0070】プロファイル重み調整プログラム 119 は、まずステップ 1101 において、正のプロファイル 120 に格納された文字列とその重みを一覧表示する。

【0071】次にステップ 1102 において、負のプロファイル 121 に格納された文字列とその重みを一覧表示する。

【0072】次にステップ 1103 において、ユーザがキーボード 101 により入力した、ユーザが重みを変更したい文字列、またはいずれかのプロファイルに追加したい文字列と、その重みを取得する。ここで、正のプロファイルにある文字列に負の重みを付与しようとした場合や、負のプロファイルにある文字列に正の重みを付与しようとした場合には、ユーザへの警告を出力するようにする等しても良い。

【0073】次にステップ 1104 において、ステップ 1103 で取得したとおりに正のプロファイル 120 または負のプロファイル 121 の内容を変更する。

【0074】以上が、プロファイル重み調整プログラム 119 の処理手順である。

【0075】ここで、図 12 にプロファイル重み調整プログラム 119 により、ユーザがプロファイルを調整する際にディスプレイ 100 に表示する入力画面の例を示す。正のプロファイル 120 の内容が 1201 に、負のプロファイル 121 の内容が 1202 に表示される。それぞれスクロールバー 1203 および 1204 を用いて、全ての内容を表示させることも可能である。ユーザがテキストボックス 1205 に重みを変更したい文字列、またはいずれかのプロファイルに追加したい文字列を入力し、重みを 1206 に入力して送信ボタン 1207 を押下する。ここで、重みを変更したい文字列文字列はテキストボックス 1205 に入力する形ではなく、表示される一覧の中からラジオボタン等により選択する形にしても良い。

【0076】次に、図 10 に示したステップ 1001 で検索条件修正制御プログラム 116 により起動される、プロファイル更新プログラム 117 の処理手順を図 13 の PAD 図を用いて説明する。

【0077】プロファイル更新プログラム 117 は、まずステップ 1301 において、ユーザがキーボード 101 により入力した文書番号と、その文書番号の文書に対するユーザの評価（「所望のものであった」あるいは「所望のものでなかった」等の評価）を読み込む。

【0078】次にステップ 1302 において、ステップ 1301 で読み込んだ文書番号に該当する文書を、テキスト 103 から登録文書保存エリア 124 に読み込む。

【0079】次にステップ 1303 において、登録文書保存エリア 124 に格納された文書から特徴文字列を抽

出し、該文書内出現回数を計数出現頻度ファイル104を参照することにより抽出し、共に特徴文字列保存エリア125に格納する。ここで、特徴文字列の抽出方法としては前掲の「特開平11-143902号広報」による方法を用いても良いし、形態素解析やニューラルネットワークによる学習データなどを用いる方法でもかまわない。

【0080】次にステップ1304において、ステップ1301で読み込んだユーザの評価が正の評価であった場合には、ステップ1306において、特徴文字列保存エリア125内の文字列の出現回数を正のプロファイルの該当文字列の重みに加算する。このとき、正のプロファイル120に無い文字列の場合には、ステップ1303で読み込んだ出現回数を重みとして付与し、該文字列を正のプロファイル120に追加する。

【0081】次にステップ1305において、ステップ1301で読み込んだユーザの評価が負の評価であった場合には、ステップ1307において、特徴文字列保存エリア125内の文字列の出現回数を負のプロファイルの該当文字列の重みから減算する。このとき、負のプロファイル121に無い文字列の場合には、ステップ1303で読み込んだ出現回数の負値を重みとして付与し、該文字列を負のプロファイル121に追加する。

【0082】ここでステップ1306、1307において重みの加減算の方法は、ユーザの評価により調整しても良い。例えばステップ1306において、ユーザが「所望のものである」という評価をした場合には、その文書内の特徴文字列の出現回数を、そのまま正のプロファイル120の該文字列の重みに足し、「やや所望のものである」という評価をした場合には、その文書内の特徴文字列の出現回数の半数を、正のプロファイル120の該文字列の重みに足す、などといった方法にしても良い。また、ステップ1306およびステップ1307で重みを加減算する特徴文字列は、ステップ1303において抽出した出現回数の上位所定数に限定しても構わない。

【0083】以上が、プロファイル更新プログラム117の処理手順である。

【0084】次に、図10に示したステップ1002において検索条件修正制御プログラム116により起動される、検索使用文字列選択プログラム118の処理手順を図14のPAD図を用いて説明する。

【0085】検索使用文字列選択プログラム118は、まずステップ1401において、総合プロファイル122の内容をクリアする。

【0086】次にステップ1402において、正のプロファイル120の中の特徴文字列のうち重みの上位所定件を抽出し、その重みと共に総合プロファイル122に追加する。

【0087】次にステップ1403において、負のプロ

ファイル121の中の特徴文字列のうち、重みの絶対値の上位所定件のもので、かつ正のプロファイル120の中の特徴文字列の重みの上位所定件に含まれないものを、総合プロファイル122に追加する。

【0088】ここでステップ1402、ステップ1403で使用する所定件数はそれぞれ異なった値でも良い。

【0089】以上が検索使用文字列選択プログラム118の処理手順である。

【0090】以上が、本実施例における各プログラムの処理手順である。

【0091】以下、本実施例において検索結果文書に対しユーザが負の評価をした場合の、検索条件の修正および再検索処理の流れを、図15を用いて説明する。

【0092】本図においては、ユーザが「高校野球」に関する文書を検索したいものとし、最初に種文書に指定した「サッカーに続き、高校野球が開幕した…」という文書1501から抽出された「サッカー」「高校」「野球」「開幕」という文字列1502が検索条件生成プログラム113により、正のプロファイル120に登録されているものとする。

【0093】ここで、「高校サッカーが開幕した…」という検索結果文書1503に対して負の評価をした場合を想定する。

【0094】まず、出現頻度ファイル104に格納された出現頻度情報のうち、ユーザが負の評価をした「高校サッカーが開幕した…」という文書1503から特徴文字列1504を抽出し、それぞれの特徴文字列の文書1503内の出現頻度とともに特徴文字列保存エリア125に読み込む。本図の例では、「高校」、「サッカー」、「開幕」、…という文字列とその出現頻度を読み込む。

【0095】次に、特徴文字列保存エリア125の文字列のうち負のプロファイル121にある文字列についてはその重みを減算し、負のプロファイル121に無い文字列については、その出現回数の負の数を重みとして負のプロファイル121に登録する。本図の例では、「高校」、「サッカー」、「開幕」、…という文字列にそれぞれ重み「-4」、「-4」、「-1」、…を付与して負のプロファイル121に追加する。

【0096】次に、正にプロファイル120の文字列のうち重みの上位所定数もの1505と、負のプロファイル121のうち重みの下位所定数1506に含まれ、かつ正のプロファイル120の文字列のうち上位所定数のもの1507に含まれないものを、総合プロファイル122に登録する。本図に示した例では、正のプロファイル120から「高校」と「野球」、負のプロファイル121から「サッカー」という文字列を選択し、総合プロファイル122に追加する。

【0097】検索時には、この総合プロファイル122の文字列とその重みにより検索を行なう。本図に示した

例では、負のプロファイル中の「高校」という文字列に関する重み値-4は検索に使用されないことになる。このことにより、「高校サッカー」の文書に負の評価をしても、「高校」という文字列の重みが下がらないため、「高校野球」よりも「プロ野球」の文書に高い類似度が算出されてしまうといった問題を防ぐことができる。

【0098】以上が、検索結果文書に対しユーザが負の評価をした場合の、検索条件の修正および再検索処理の流れである。

【0099】以上示したように本実施例によれば、ユーザが「所望のものでない」と評価した文書から抽出された文字列のうち、ユーザが「所望のものである」と評価した文書から抽出された文字列を、重みを下げる対象から除外する形態をとる。そのため、ユーザの所望ではない概念を表す文字列のみの重みを適切に減算することができる。したがって、ユーザが「所望のものでない」と評価した文書から抽出した文字列の重みを単純に減算すると、ユーザの所望の概念を表す文字列の重みまで減算してしまい、検索結果が改善しない、といった問題を解決できる。

【0100】また、本実施例によれば、検索結果文書の内容を表示する際、検索条件データに保存されている文字列の重み正負により文字列を別の形式でハイライト表示する形態をとる。

【0101】この方法により、ユーザは、検索結果文書がどの程度所望の内容を示しているかを視覚的に容易に判断できる。また、正の重みが付与された文字列や負の重みが付与された文字列として、どのようなものが所望文書やノイズ文書に含まれているかを見ることにより、次回以降のプロファイルの調整に役立てることができるようになる。

【0102】また、本実施例によれば、検索条件データの中の文字列のうち検索に用いる文字列をユーザが選択、あるいはそれぞれの文字列の重みをユーザが調整する形態をとる。

【0103】この方法により、ユーザの所望する内容を特徴付けるものでないものを、検索に使用することを防ぐことができ、適切な検索結果を得られるようになる。

【0104】図13に示したプロファイル更新プログラムの処理においては、ユーザが負の評価をした際に、評価対象文書から抽出した文字列を負のプロファイル121に追加した後、総合プロファイル122に追加する文字列を選択する形態をとっている。ここで図16に示すように、評価対象文書から抽出した文字列のうち、負のプロファイル121に追加する文字列を選択する形態をとっても良い。

【0105】すなわち、図16のステップ1305において、ステップ1301で読み込んだユーザの評価が負の評価であった場合には、ステップ1307を実行する前に図16に示すプロファイル更新用文字列選択ステッ

プ1601を実行しても良い。ここでプロファイル更新用文字列選択ステップ1601は、特徴文字列保存エリア125の文字列のうち、正のプロファイル120中の重みの上位のものに含まれるものを、特徴文字列保存エリア125からクリアするステップである。これにより、正のプロファイル120に追加されているユーザの所望の概念を表す文字列に、負の重みを付与し負のプロファイル121に追加してしまうことを防ぐことができる。

【0106】以下、本発明の第二の実施例について説明する。

【0107】第一の実施例においては、検索時に使用する文字列、または検索条件の修正時にプロファイルに追加する文字列をシステムが自動的に選択する。したがって、検索結果文書に対するユーザの評価が不適切な場合には、検索精度が向上しないという問題がある。

【0108】以上の問題を解決するために、本発明の第二の実施例では、ユーザが正または負の評価をした文書から抽出される文字列を一覧表示し、正の重みまたは負の重みを付与する文字列をユーザが選択する手段を提供するものである。

【0109】本実施例は図1に示す第一の実施例とほぼ同様の構成をとる。ここで図17に示すように検索条件修正制御プログラム116aはプロファイル更新用文字列ユーザ選択プログラム1701、プロファイル更新プログラム117a、および検索使用文字列選択プログラム118により構成される。また、図18に示すようにプロファイル更新プログラム117aの処理手順が、第一の実施例におけるプロファイル更新プログラム117と異なる。

【0110】以下、第二の実施例における、プロファイル更新プログラム117aの処理手順について図18のPAD図を用いて説明する。

【0111】まずプロファイル更新プログラム117aは、まずステップ1801において、ユーザがキーボード101により入力した文書番号と、その文書番号の文書に対するユーザの評価（「所望のものであった」あるいは「所望のものでなかった」等の評価）を読み込む。

【0112】次にステップ1802において、ステップ1801で読み込んだ文書番号に該当する文書を、テキスト103から登録文書保存エリア124に読み込む。

【0113】次にステップ1803において、登録文書保存エリア124に格納された文書から特徴文字列を抽出し、該文書内出現回数を計数出現頻度ファイル104を参照することにより抽出し、共に特徴文字列保存エリア125に格納する。ここで、特徴文字列の抽出方法としては前掲の「特開平11-143902号広報」による方法を用いても良いし、形態素解析やニューラルネットワークによる学習データなどを用いる方法でもかまわない。

【0114】次にステップ1804において、プロフィール更新用文字列ユーザ選択プログラム1701を起動し、ステップ1803において読み込んだ文字列のうちユーザが選択しなかった文字列を、特徴文字列保存エリア125からクリアする。

【0115】次にステップ1805において、ステップ1801で読み込んだユーザの評価が正の評価であった場合には、ステップ1807において、特徴文字列保存エリア125の文字列の出現回数を正のプロファイルの該当文字列の重みに加算する。このとき、正のプロファイル120に無い文字列の場合には、ステップ1803で読み込んだ出現回数を重みとして付与し、該文字列を正のプロファイル120に追加する。

【0116】次にステップ1806において、ステップ1801で読み込んだユーザの評価が負の評価であった場合には、ステップ1808において、特徴文字列保存エリア125の文字列の出現回数を負のプロファイルの該当文字列の重みから減算する。このとき、負のプロファイル121に無い文字列の場合には、ステップ1803で読み込んだ出現回数の負値を重みとして付与し、該文字列を負のプロファイル121に追加する。

【0117】ここでステップ1807、1808において重みの加減算の方法は、ユーザの評価により調整しても良い。例えばステップ1807において、ユーザが「所望のものである」という評価をした場合には、その文書内の特徴文字列の出現回数を、そのまま正のプロファイル120の該文字列の重みに足し、「やや所望のものである」という評価をした場合には、その文書内の特徴文字列の出現回数の半数を、正のプロファイル120の該文字列の重みに足す、などといった方法にしても良い。また、ステップ1807およびステップ1808で重みを加減算する特徴文字列は、ステップ1803において抽出した出現回数の上位所定数に限定しても構わない。

【0118】以上が、プロフィール更新プログラム117aの処理手順である。

【0119】次に図18に示したステップ1804でプロフィール更新プログラム117aにより起動される、プロフィール更新用文字列ユーザ選択プログラム1701の処理手順を、図19のPAD図を用いて説明する。

【0120】まずステップ1901において、特徴文字列保存エリア125内の特徴文字列を一覧表示する。

【0121】次にステップ1902において、ステップ1901で表示した文字列のうち、ユーザが選択しなかった文字列を取得し、該文字列の情報を特徴文字列保存エリア125からクリアする。

【0122】以上がプロフィール更新用文字列ユーザ選択プログラム1701の処理手順である。

【0123】ここで、プロフィール更新用文字列ユーザ選択プログラム1701により、ユーザがプロフィール

に追加したい文字列を選択する画面の例を図20に示す。ウィンドウ2001に、ユーザが評価した文書から抽出される特徴文字列がチェックボックスと共に表示される。特徴文字列が多数ある場合はスクロールバー2002を用いてすべての文字列をウィンドウ2001内で参照することができる。ユーザは、ウィンドウ2001内の文字列のうち、プロフィールに追加したい文字列のチェックボックスをチェックし、送信ボタン2003を押下する。

【0124】なお、文字列の選択方法は図20の例のようにチェックボックスを用いたものでも良いし、各文字列に識別番号を付与して識別番号と共に一覧表示するようにし、文字列の識別番号により選択する方法でも良い。

【0125】以下、本実施例において検索結果テキストに対しユーザが負の評価をした場合の、検索条件の修正および再検索処理の流れを、図21を用いて説明する。

【0126】本図においては、ユーザが「高校野球」に関するテキストを検索したいものとし、最初に種文書に指定した「サッカーに続き、高校野球が開幕した…」というテキスト2101から抽出されたサッカー「高校」「野球」「開幕」という文字列2102が検索条件生成プログラム113により、正のプロファイル120に登録されているものとする。

【0127】ここで、「高校サッカーの1回戦が…」という検索結果テキストに対して負の評価をした場合を想定する。

【0128】まず、出現頻度ファイル104に格納された出現頻度情報のうち、ユーザが負の評価をした「高校サッカーの1回戦が…」という文書2103から特徴文字列2104を抽出し、それぞれの特徴文字列の文書2103内の出現頻度とともに特徴文字列保存エリア125に読み込む。本図の例では、「高校」、「サッカー」、「1回戦」、…という文字列とその出現頻度が読み込まれる。

【0129】次に、前述した図20の画面でユーザが選択した文字列の情報を、文字列保存エリア125からクリアする。本図の例では、ユーザが「高校野球」に関するテキストを所望しており、「サッカー」に関するテキストは所望ではない。したがってユーザは「サッカー」という文字列のみに負の重みを加えると指定するものとする。このとき、文字列保存エリア125から、「高校」および「1回戦」という文字列とその重みをクリアする。

【0130】次に、出現頻度情報2104のうち負のプロファイル121にある文字列についてはその重みを減算し、負のプロファイル121に無い文字列については、その出現回数の負の数を重みとして負のプロファイル121に登録する。本図の例では、「サッカー」という文字列に重み「-4」を付与して正のプロファイル1

20に追加する。

【0131】次に、正にプロファイル120の文字列のうち重みの上位所定数もの2105と、負のプロファイル121のうち重みの下位所定数2106に含まれ、かつ正のプロファイル120の文字列のうち上位所定数もの2107に含まれないものを、総合プロファイル122に登録する。検索時には、この総合プロファイル122の文字列とその重みにより検索を行なう。

【0132】以上のように、本図に示した例では、「高校サッカーの1回戦が…」というテキストに負の評価をしても、「高校」という文字列の重みが下がらないため、「高校野球」よりも「プロ野球」のテキストに高い類似度が算出されてしまうといった問題を防ぐことができる。また、正のプロファイル120に無い「1回戦」という文字列の重みがさがらないため、「高校野球の1回戦」といったユーザが所望するテキストの類似度が下がってしまうといった問題を防ぐことができる。

【0133】以上が、検索結果テキストに対しユーザが負の評価をした場合の、検索条件の修正および再検索処理の流れである。

【0134】なお、本実施例において検索結果文書に対しユーザが正の評価をした場合にも同様に、正のプロファイルに追加する文字列を選択することができる。したがって、正の評価をした文書から抽出されるがユーザの概念を表す文字列ではない文字列に、正の重みを付与してしまうことを防ぐことができる。

【0135】以上が、本発明の第二の実施例である。

【0136】以上示したように本実施例によれば、ユーザが「所望のものでない」と評価した文書から抽出された文字列のうち、ユーザが所望する概念を表す文字列をユーザが指定することにより、該文字列を重みを下げる対象から除外する形態をとる。そのため、ユーザの所望ではない概念を表す文字列のみの重みを適切に減算することができる。したがって、ユーザが「所望のものでない」と評価した文書から抽出した文字列の重みを単純に減算すると、ユーザの所望の概念を表す文字列の重みまで減算してしまい、検索結果が改善しない、といった問題を解決できる。

【0137】また、ユーザが「所望のものである」と評価した文書から抽出された文字列のうち、ユーザが所望する概念を表さない文字列をユーザが指定することにより、該文字列を重みを上げる対象から除外する形態をとる。そのため、ユーザの所望する概念を表す文字列のみの重みを適切に加算することができる。したがって、ユーザが「所望のものである」と評価した文書から抽出した文字列の重みを単純に加算すると、ユーザの所望の概念を表さない文字列の重みまで加算してしまい、検索結果が改善しない、といった問題を解決できる。

【0138】なお、第一、第二の実施例において、ひとつの検索結果文書に対しユーザが評価を入力し、その評

価を反映した検索結果を出力するようにしたが、複数の検索結果文書に対しそれぞれ異なった評価を一度に入力し、それらの評価を反映した検索結果を出力するようにしても構わない。

【0139】また、第一、第二の実施例において、最初に種文書を設定し、その種文書に類似した内容を持つ文書を検索するものとしたが、最初にキーワードを設定する全文検索を行なう形式にしても良い。その場合には、図7に示した検索条件生成プログラム113のステップ702、703のかわりに、入力したキーワードを所定の重みを付与して正のプロファイル120、および総合プロファイル122に追加すれば良い。

【0140】本実施例によれば、ユーザの所望の概念を表す単語の重みを減算しないため、ユーザが「所望のものでない」といった評価を与えた検索結果文書から抽出した情報をもとに検索結果を改善することができる。

【0141】

【発明の効果】本発明によれば、ユーザが「所望のものでない」といった評価を与えた文書から抽出した情報のうち適切なものを使用して、検索結果を改善することができる。

【図面の簡単な説明】

【図1】本発明の第一の実施例の構成を示す図である。

【図2】従来技術によるレリバンスフィードバック処理の例を示す図である。

【図3】従来技術によるレリバンスフィードバック処理により検索結果が改善しない例を示す図である。

【図4】本発明の第一の実施例におけるシステム制御プログラム110の処理手順を示すPAD図である。

【図5】本発明の第一の実施例における文書登録プログラム111の処理手順を示すPAD図である。

【図6】本発明の第一の実施例における検索制御プログラム112の処理手順を示すPAD図である。

【図7】本発明の第一の実施例における検索条件生成プログラム113の処理手順を示すPAD図である。

【図8】本発明の第一の実施例における類似文書検索プログラム114の処理手順を示すPAD図である。

【図9】本発明の第一の実施例における検索結果文書内容表示プログラム115の処理手順を示すPAD図である。

【図10】本発明の第一の実施例における検索条件修正制御プログラム116の処理手順を示すPAD図である。

【図11】本発明の第一の実施例におけるプロファイル重み調整プログラム119の処理手順を示すPAD図である。

【図12】本発明の第一の実施例において、ユーザがプロファイルを調整する際にディスプレイ100に表示する入力画面の例を示す図である。

【図13】本発明の第一の実施例におけるプロファイル

更新プログラム117の処理手順を示すPAD図である。

【図14】本発明の第一の実施例における検索使用文字列選択プログラム118の処理手順を示すPAD図である。

【図15】本発明の第一の実施例において、検索結果文書に対しユーザが負の評価をした場合の、検索条件の修正および再検索処理の流れを示す図である。

【図16】本発明の第一の実施例におけるプロファイル更新プログラムの処理117の処理の一形態を示すPAD図である。

【図17】本発明の第二の実施例における検索条件修正プログラム116aの構成を示すPAD図である。

【図18】本発明の第二の実施例におけるプロファイル更新プログラム117aの処理手順を示すPAD図である。

【図19】本発明の第二の実施例におけるプロファイル更新用文字列ユーザ選択プログラム1701の処理手順を示すPAD図である。

【図20】本発明の第二の実施例において、ユーザがプロファイルに追加したい文字列を選択する画面の例を示すPAD図である。

【図21】本発明の第二の実施例において、検索結果文書に対しユーザが負の評価をした場合の、検索条件の修正および再検索処理の流れを示す図である。

【符号の説明】

100 ディスプレイ

101 キーボード

102 中央演算処理装置(CPU)

103 テキスト

104 出現頻度ファイル

105 磁気ディスク装置

106 フロッピディスクドライブ(FDD)

107 フロッピディスク

108 バス

109 主メモリ

110 システム制御プログラム

111 文書登録プログラム

112 検索制御プログラム

113 検索条件生成プログラム

114 類似文書検索プログラム

115 検索結果文書内容表示プログラム

116 検索条件修正制御プログラム

117 プロファイル更新プログラム

118 検索使用文字列選択プログラム

119 プロファイル重み調整プログラム

120 正のプロファイル

121 負のプロファイル

122 総合プロファイル

123 種文書保存エリア

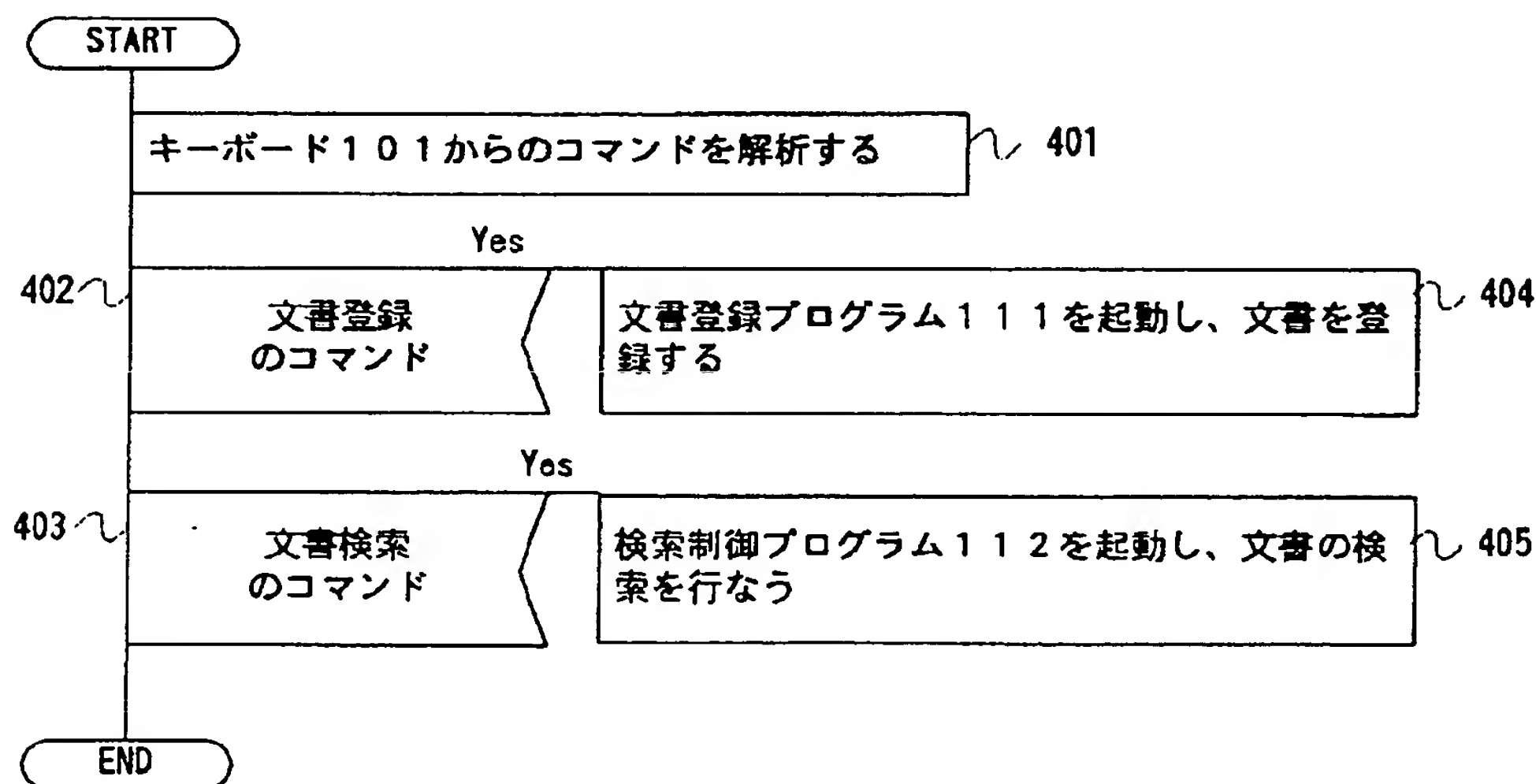
124 登録文書保存エリア

125 特徴文字列保存エリア

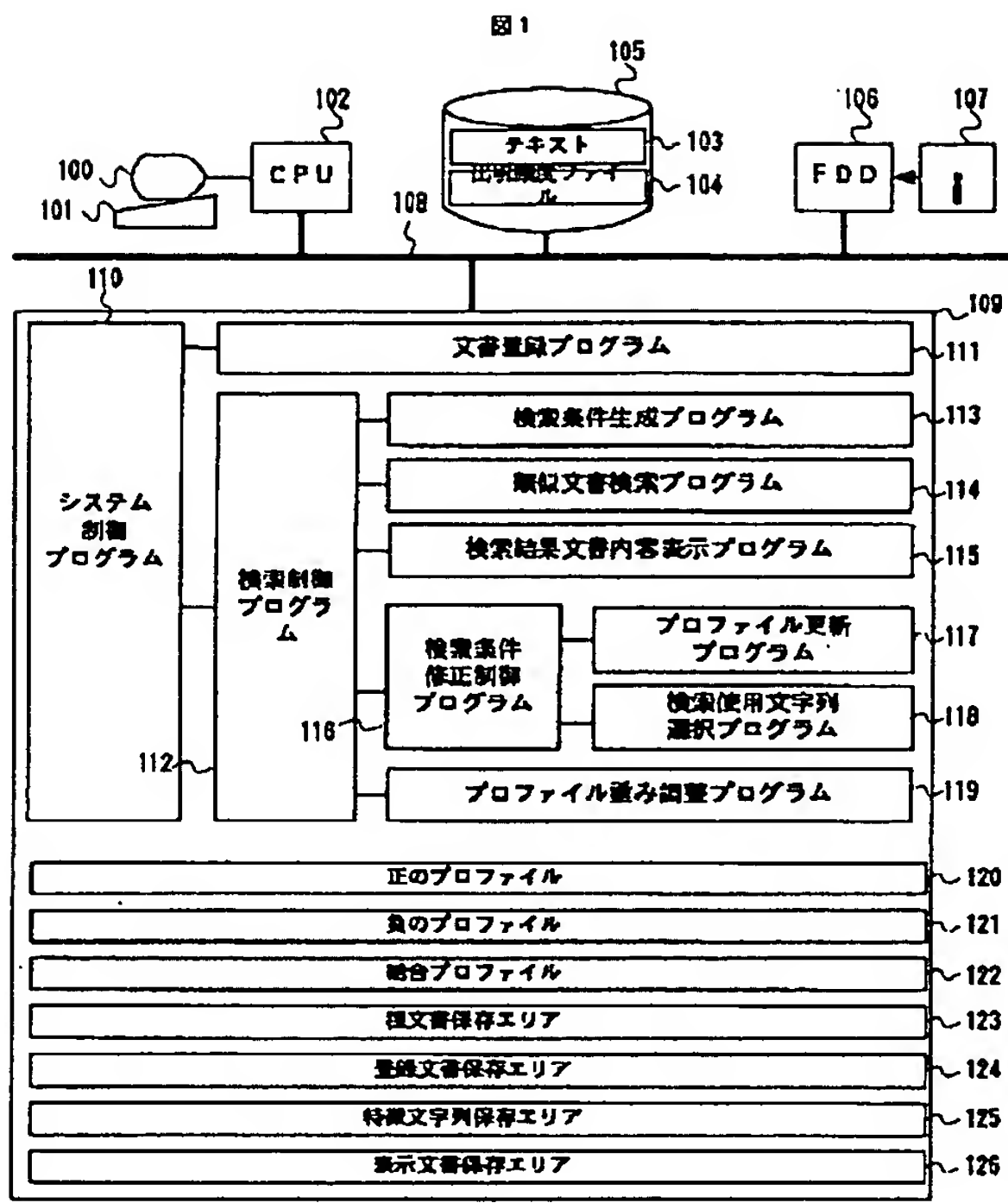
126 表示文書保存エリア

【図4】

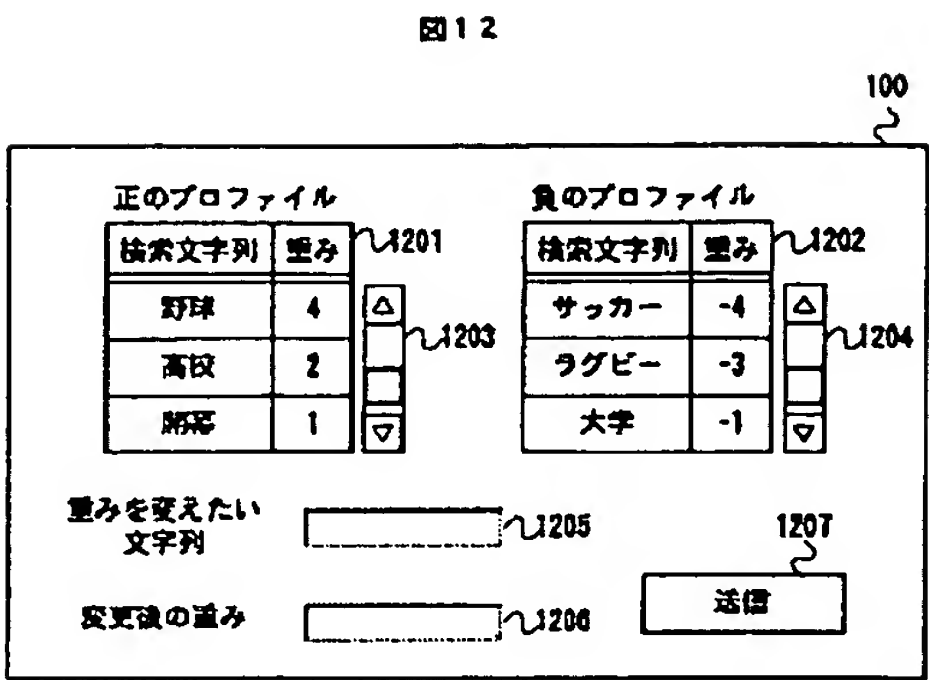
図4



【図 1】

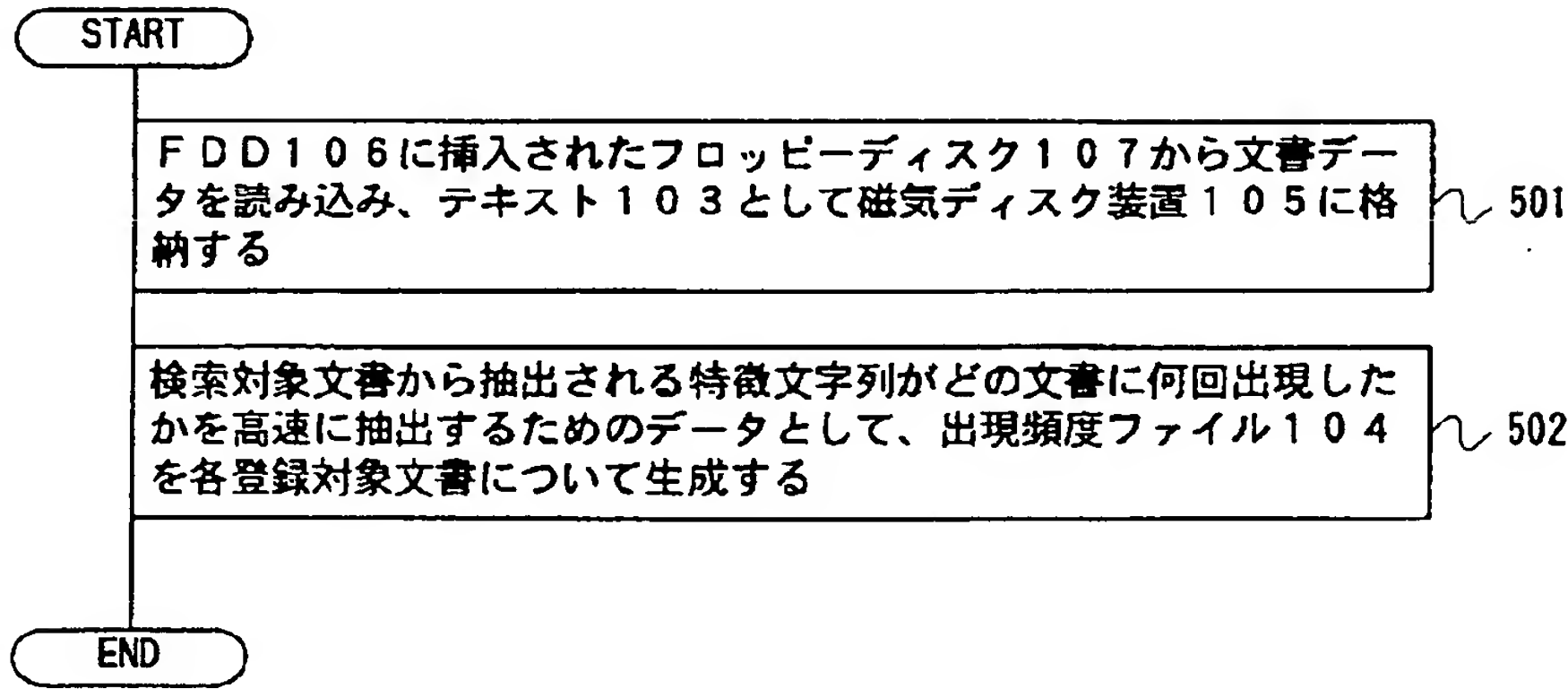


【図 1 2】



【図 5】

図 5



【図2】

図2

種文書：

サッカーに続き、高校野球が
開幕した…

単語抽出

単語と出現回数

サッカー 1
高校 2
野球 4
開幕 1

検索条件
を生成

種文書設定直後
の検索条件

| 検索文字列 | 重み |
|-------|----|
| 野球 | 4 |
| 高校 | 2 |
| サッカー | 1 |
| 開幕 | 1 |
| ⋮ | ⋮ |



ノイズ文書：

サッカーのW杯は…

単語抽出

単語と出現回数

サッカー 4
W杯 2

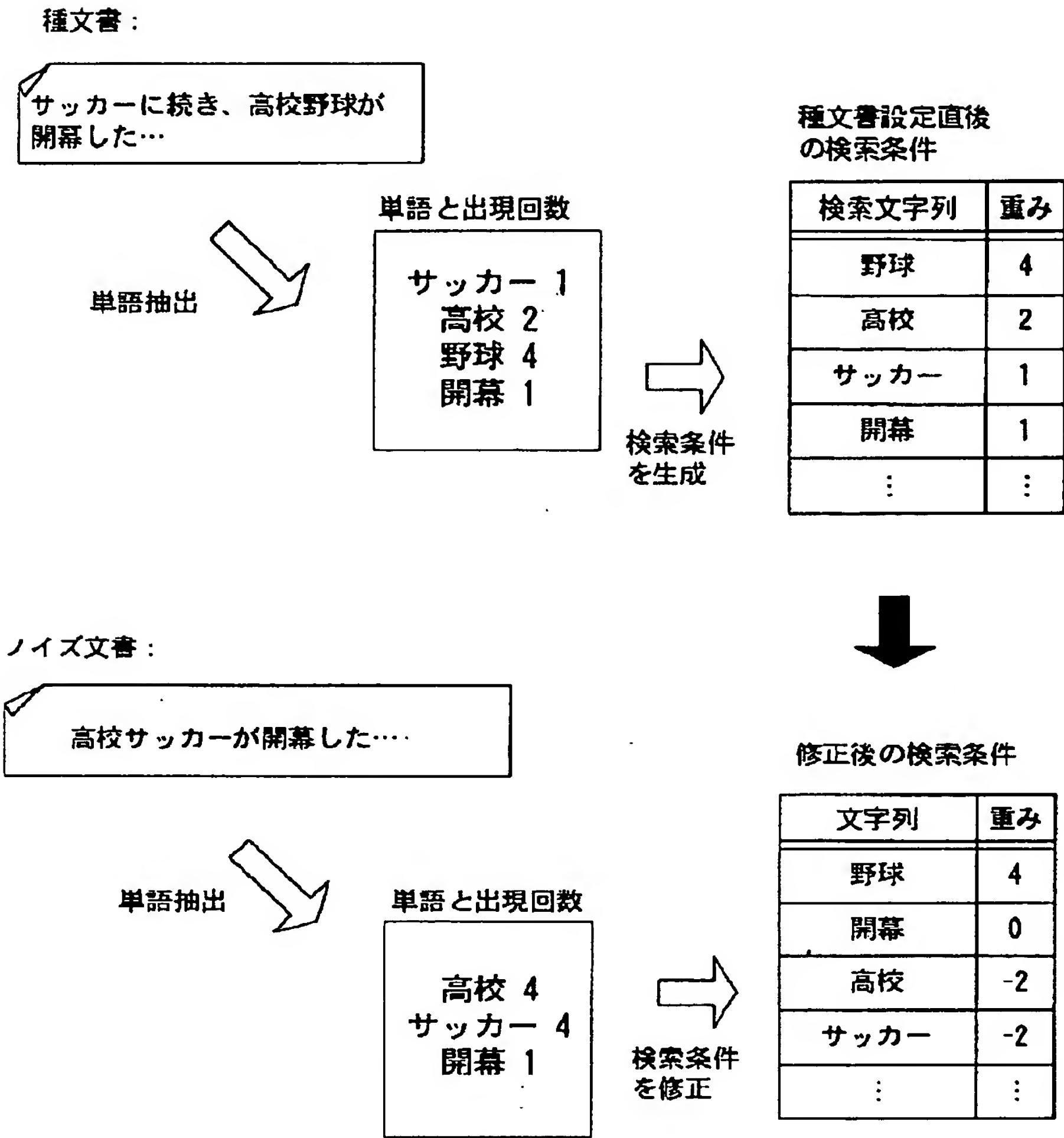
検索条件
を修正

修正後の検索条件

| 文字列 | 重み |
|------|----|
| 野球 | 4 |
| 高校 | 2 |
| 開幕 | 1 |
| W杯 | -2 |
| サッカー | -3 |
| ⋮ | ⋮ |

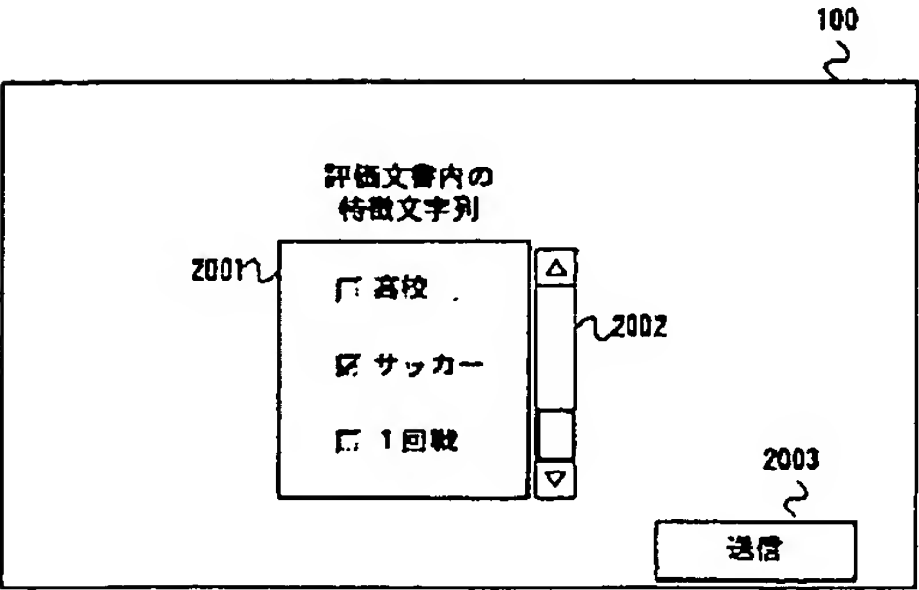
【図3】

図3



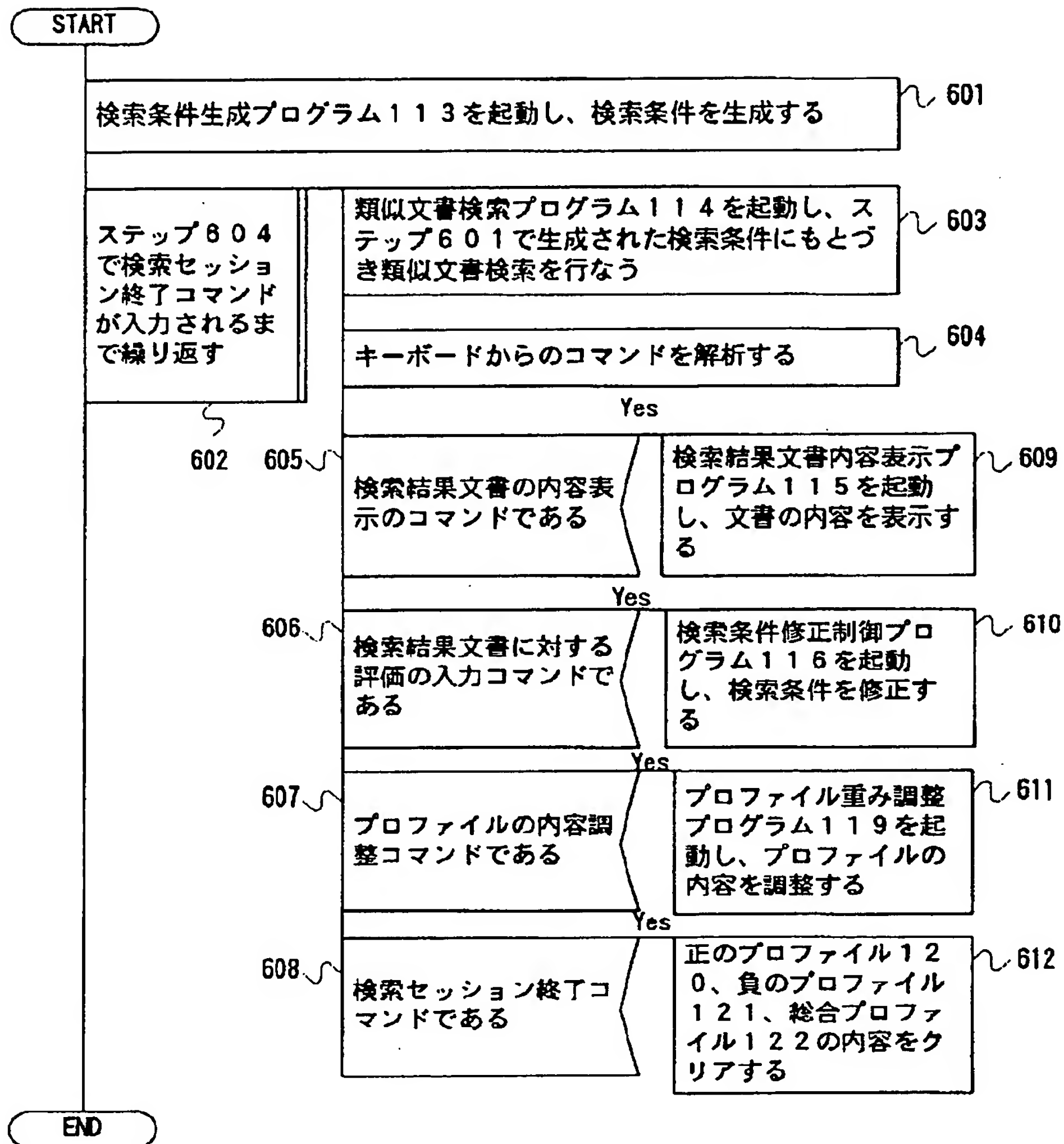
【図20】

図20



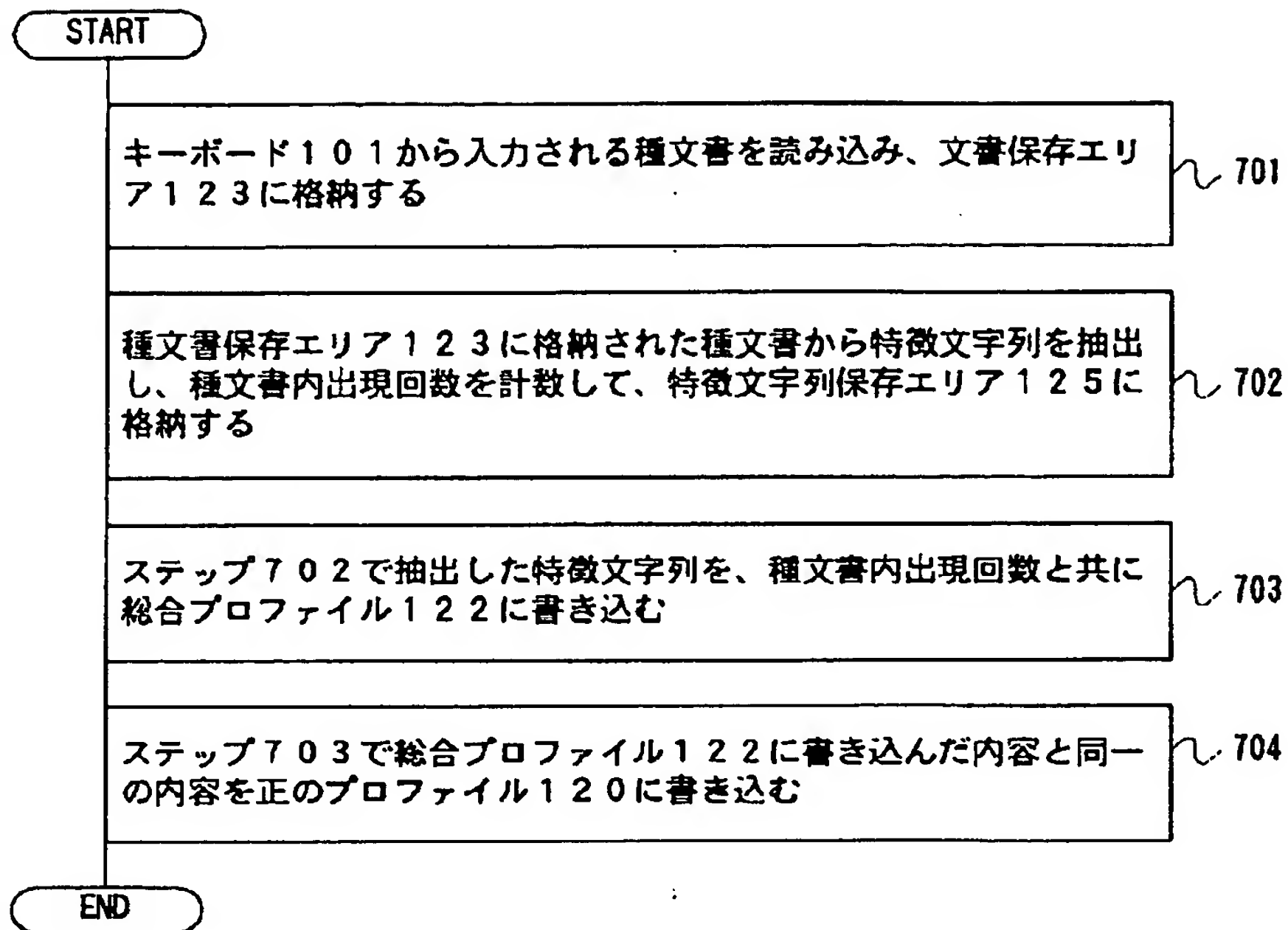
【図6】

図6



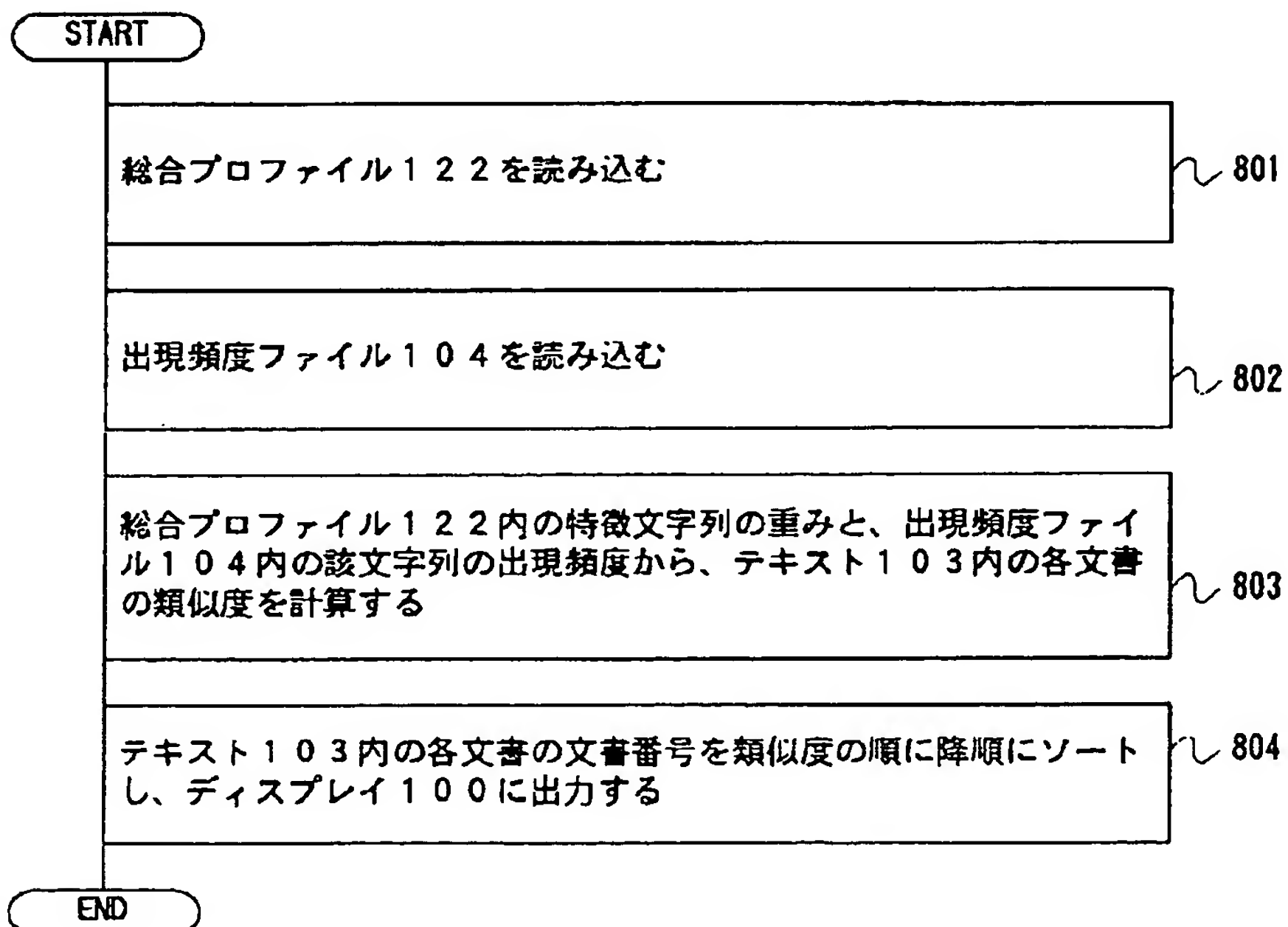
【図7】

図7



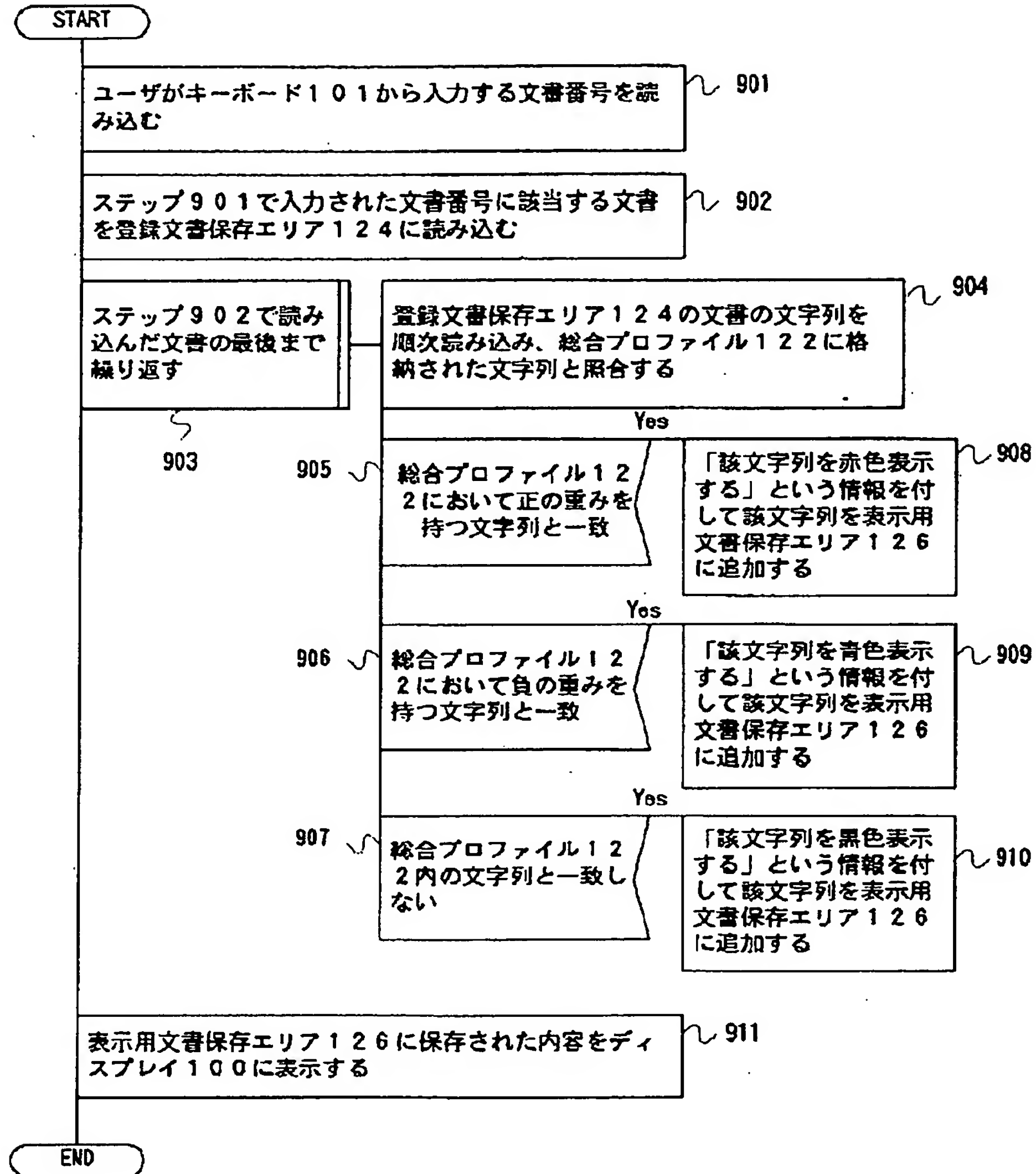
【図8】

図8



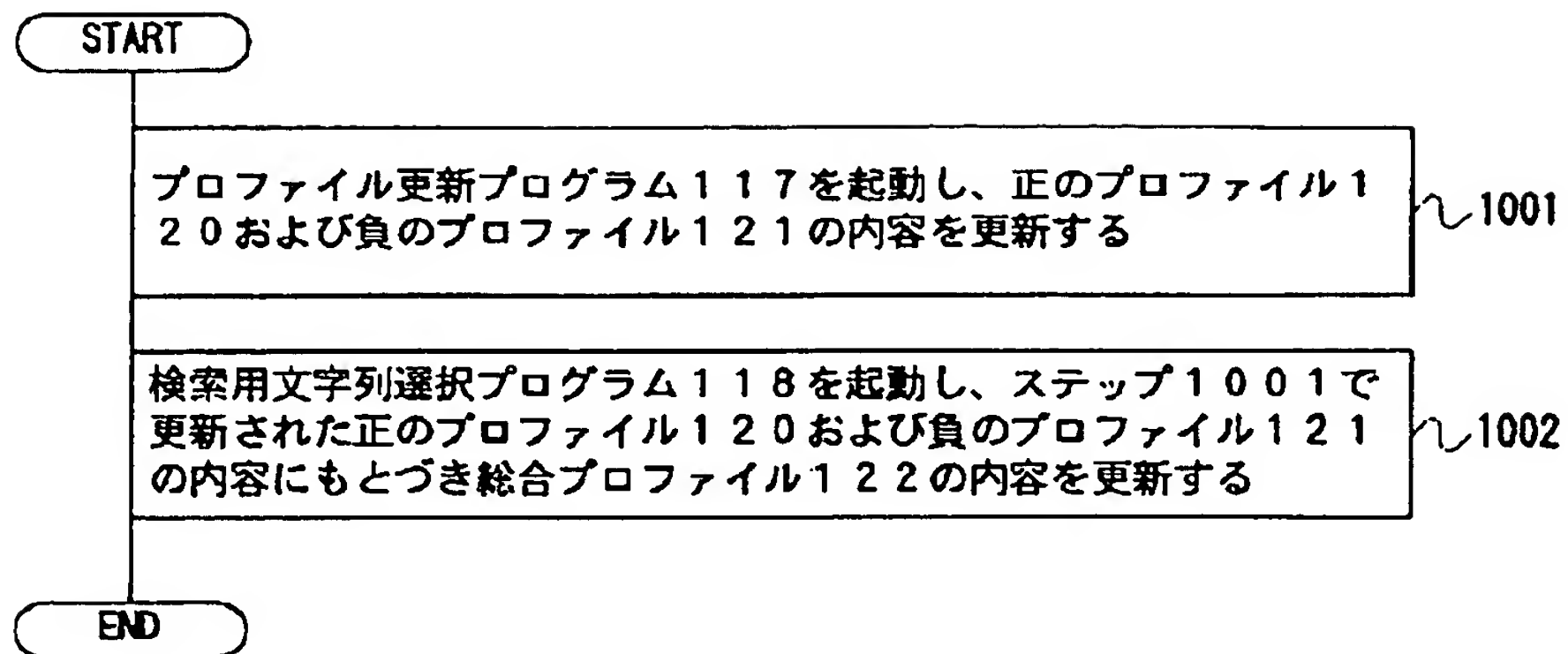
【図9】

図9



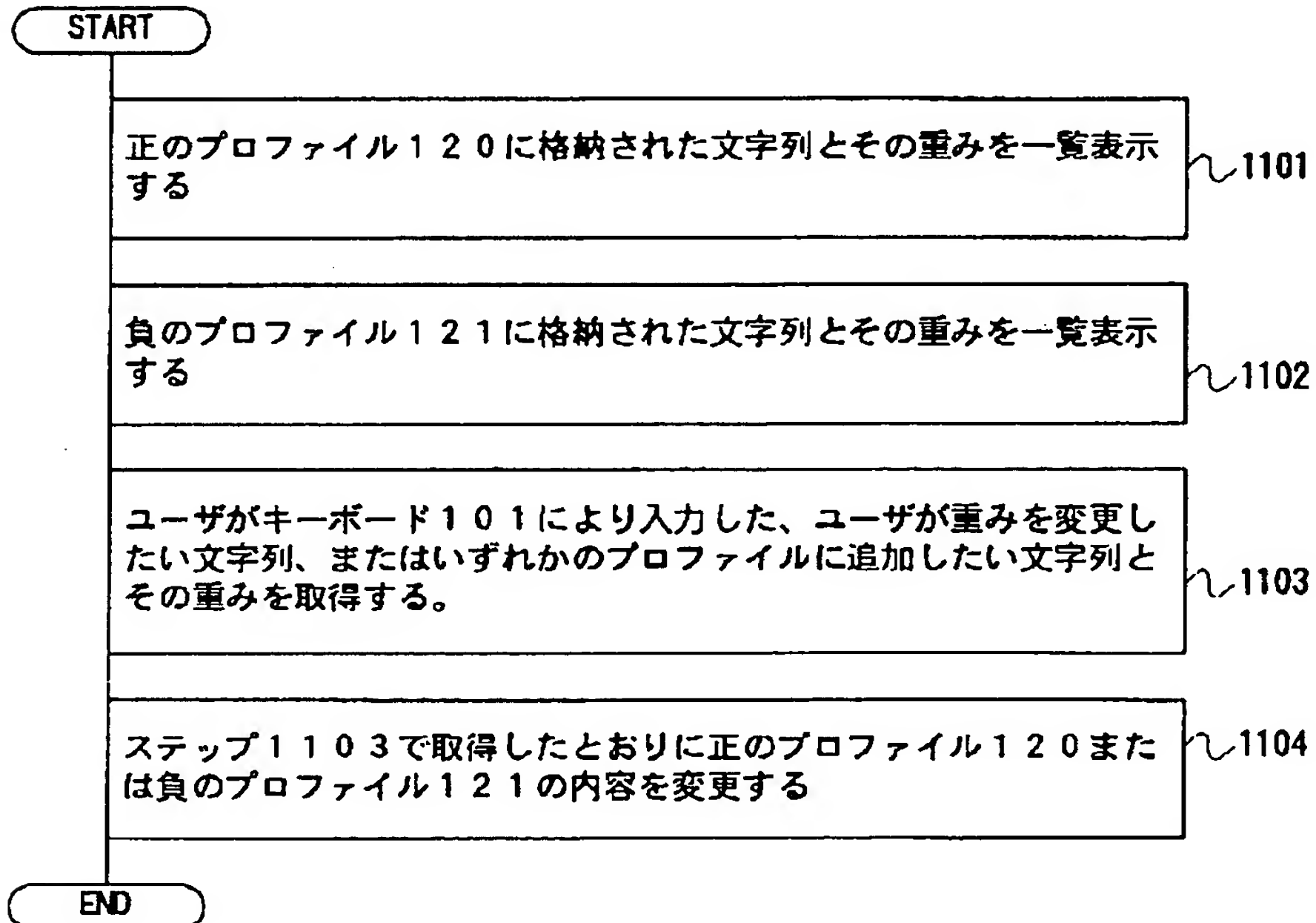
【図 10】

図 10



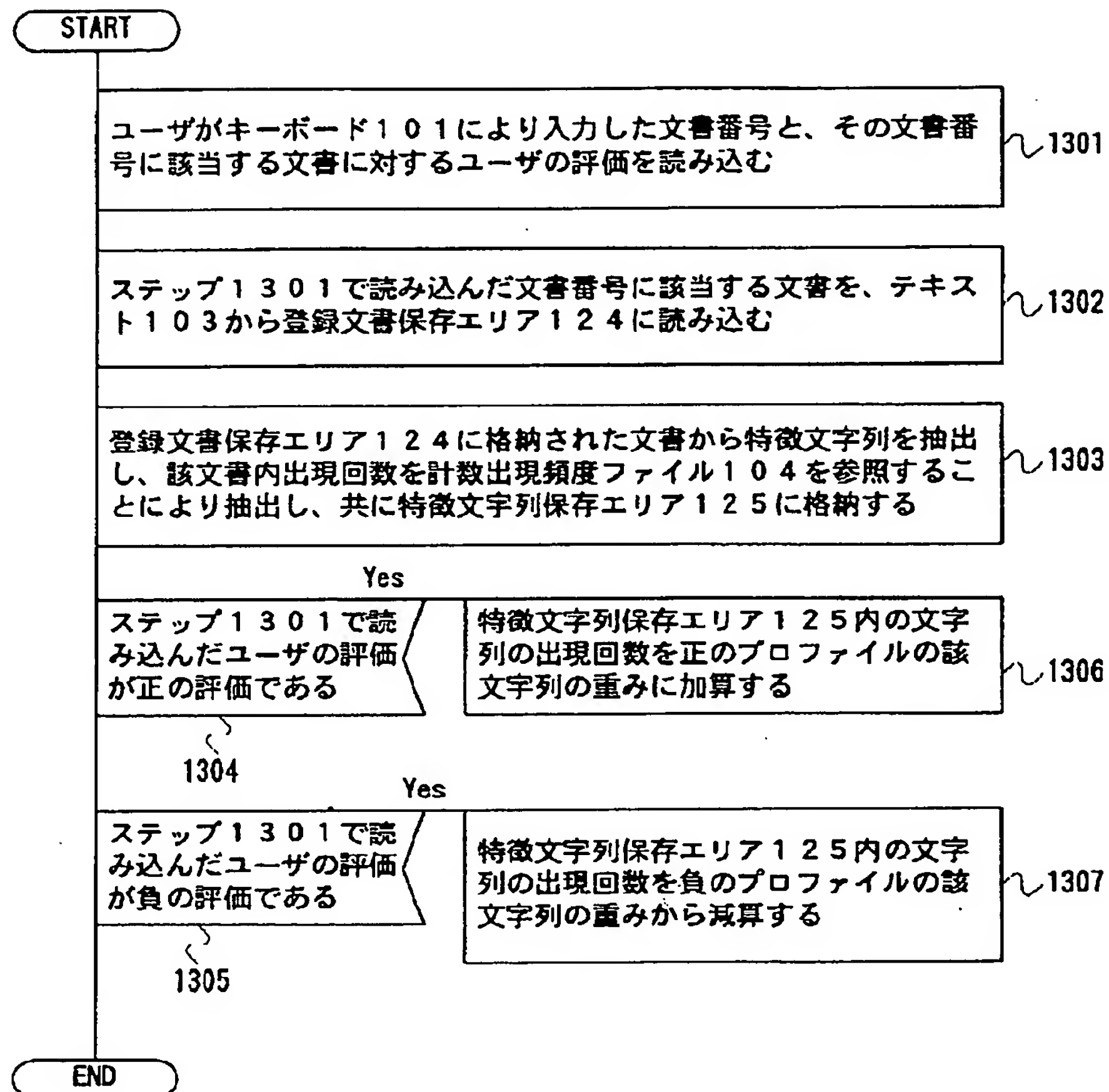
【図 11】

図 11



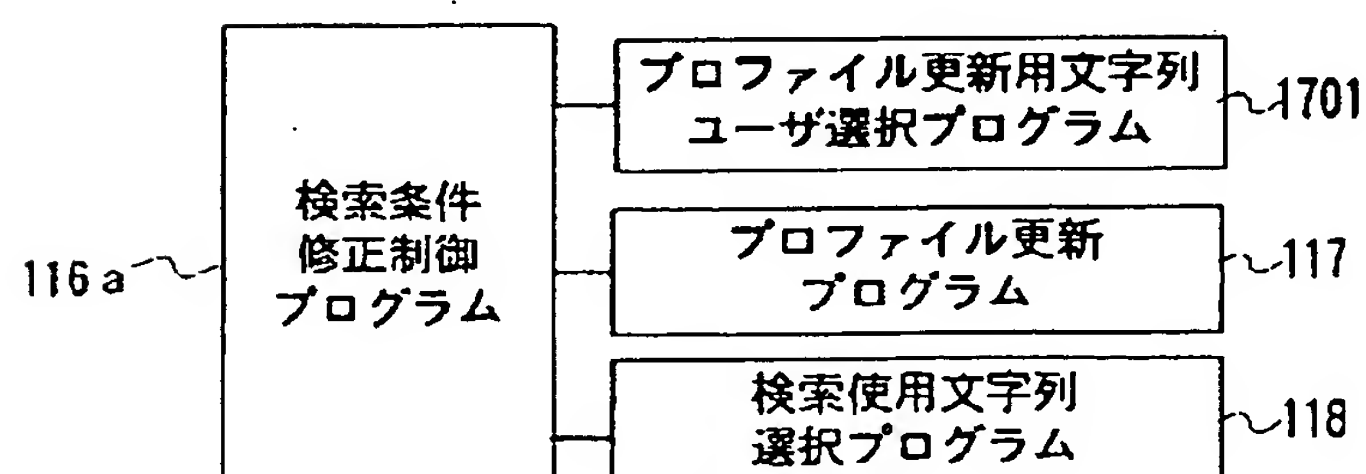
【図13】

図13



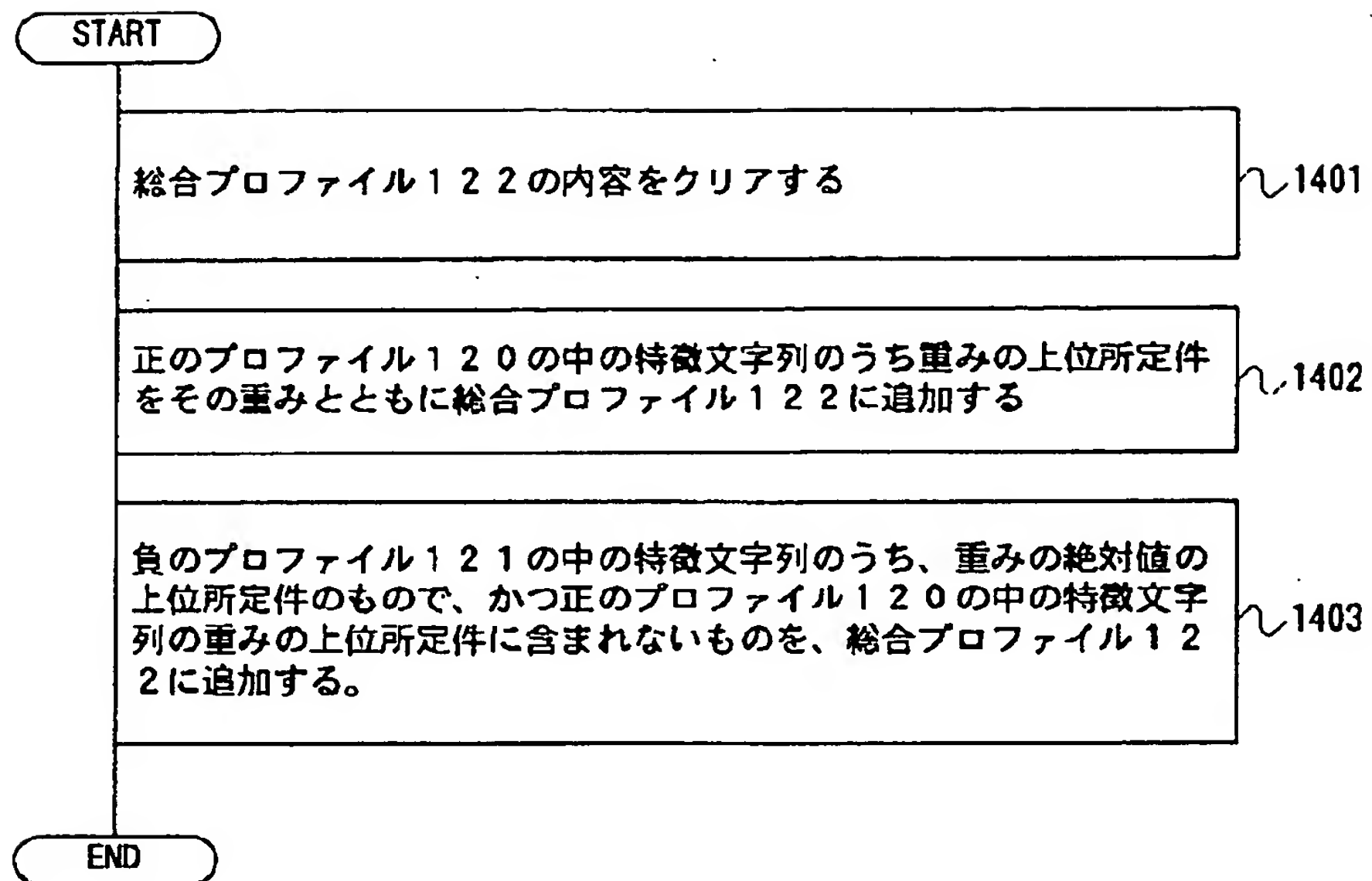
【図17】

図17



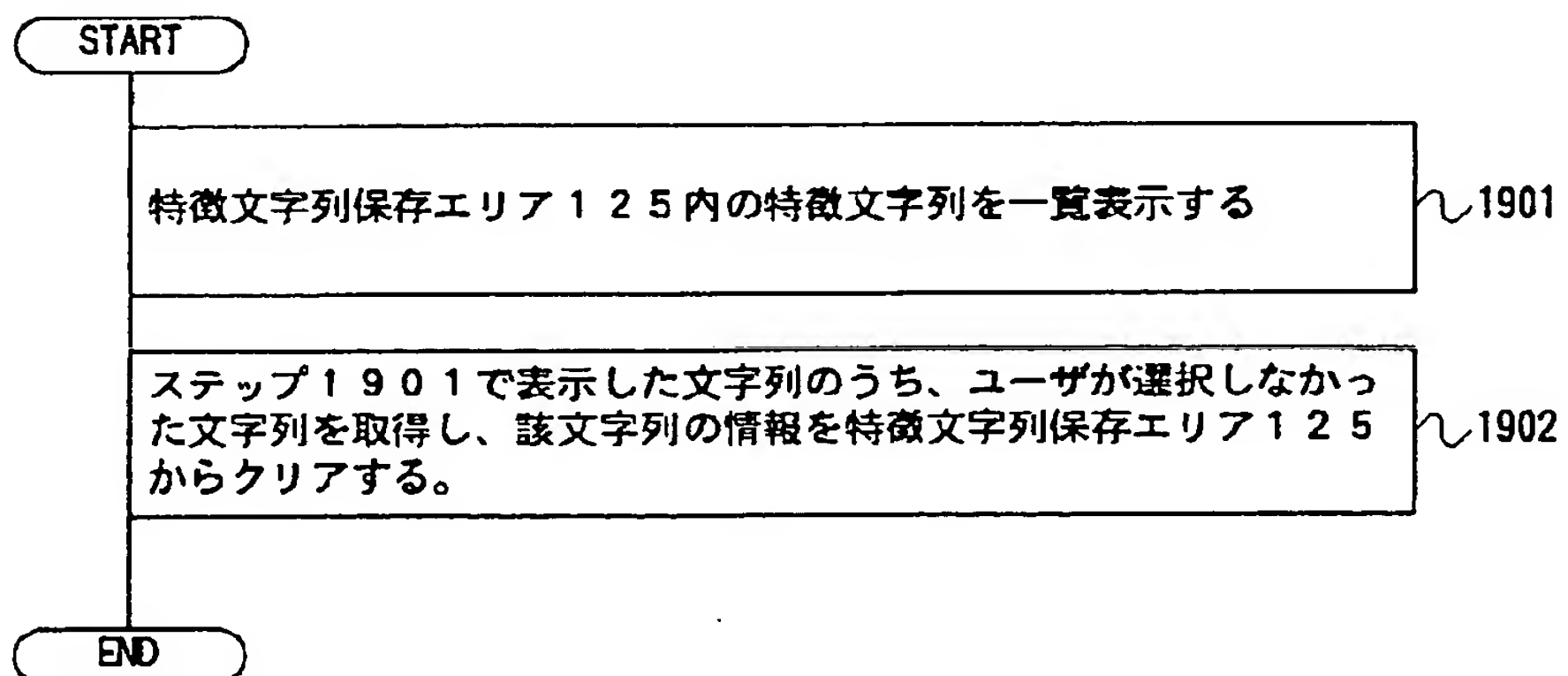
【図 14】

図 14



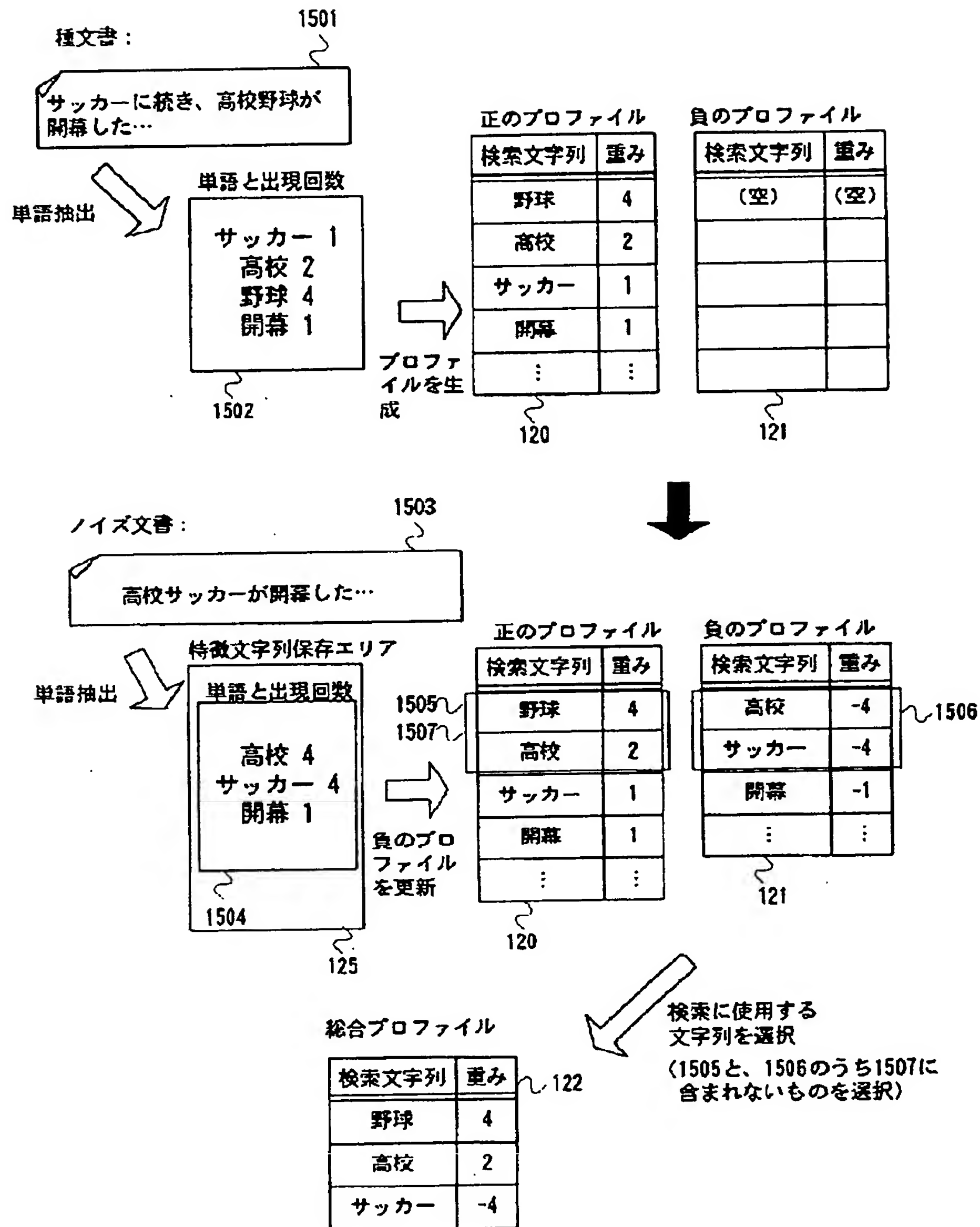
【図 19】

図 19



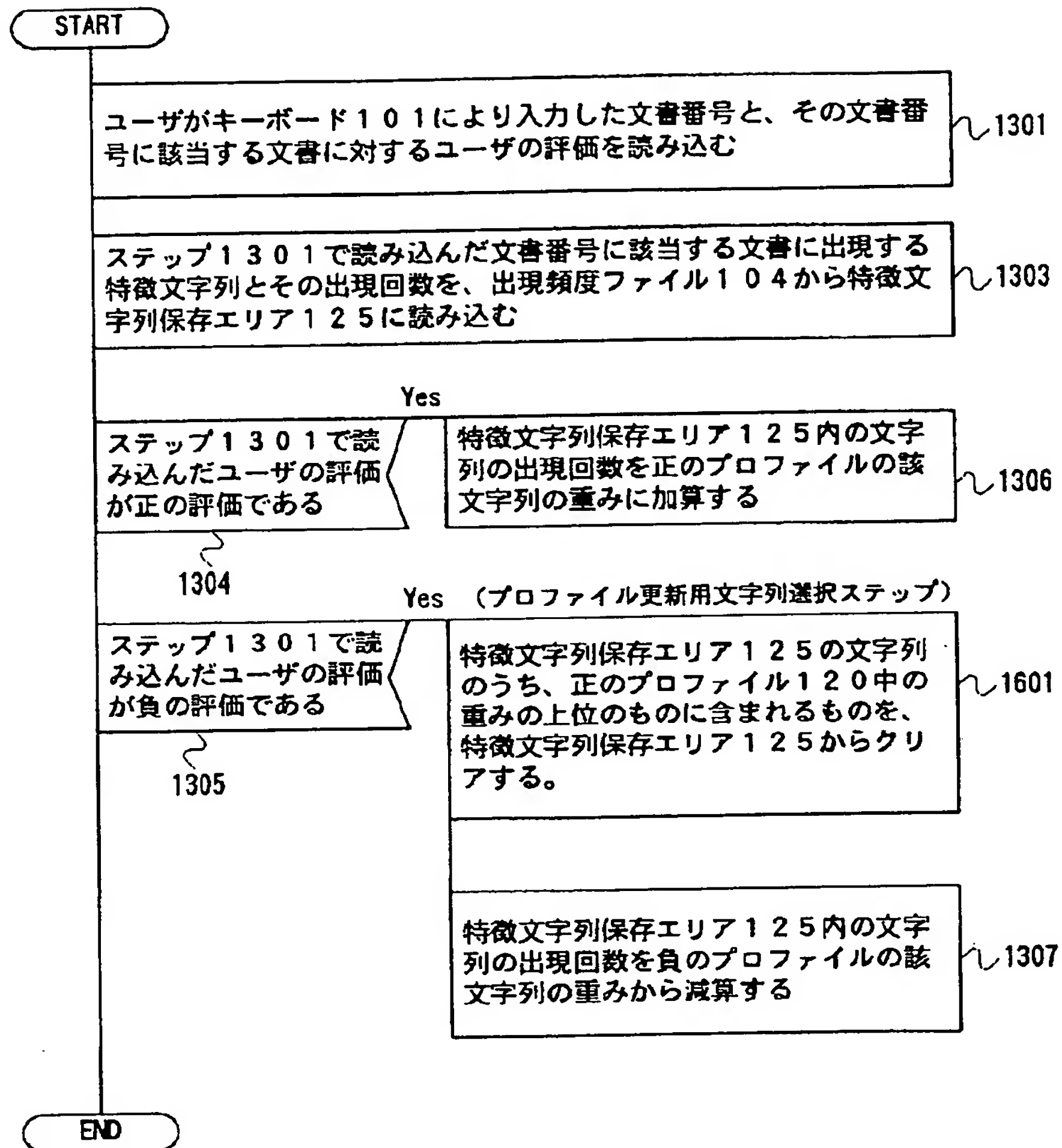
【図15】

図15



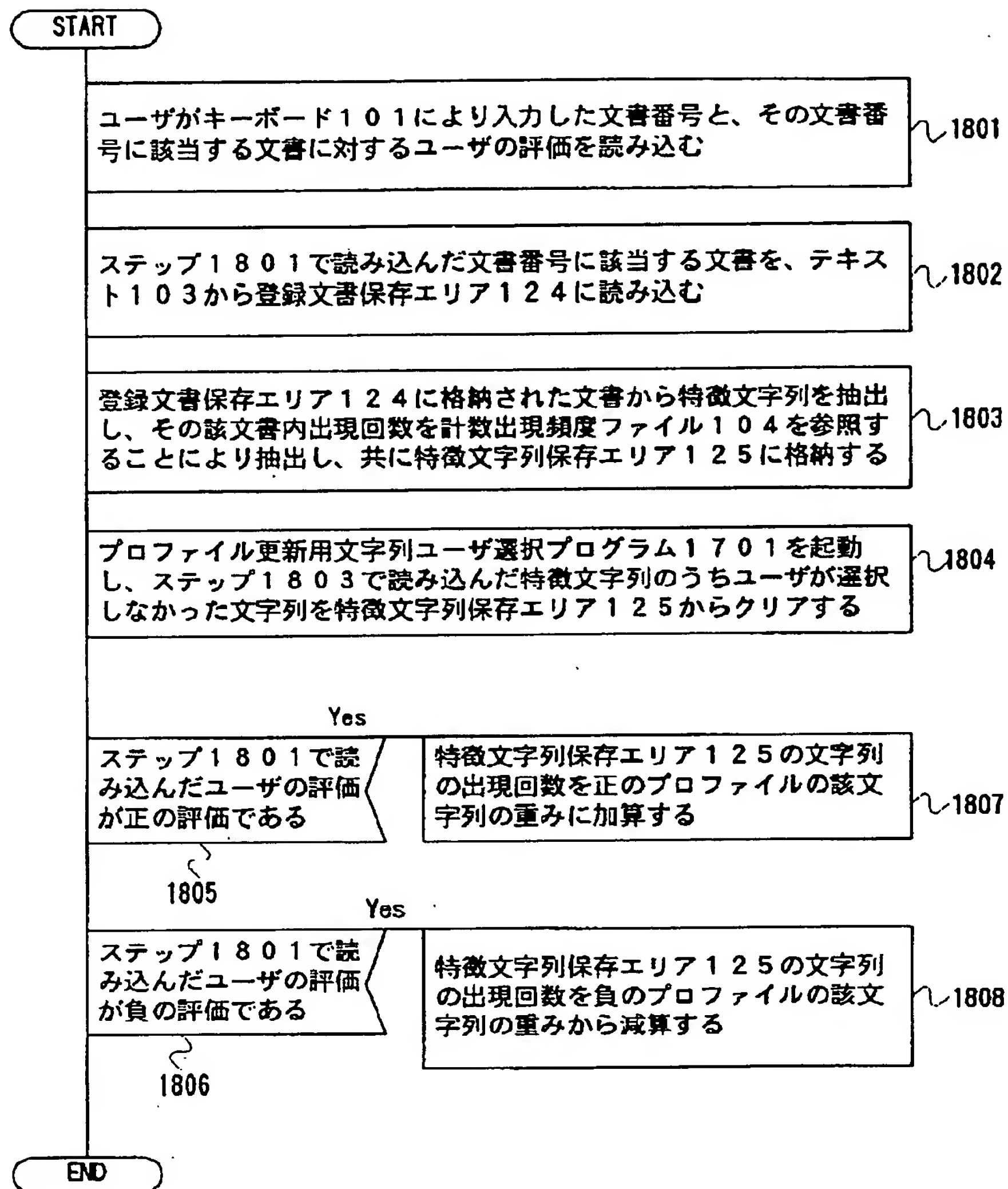
【図 16】

図 16



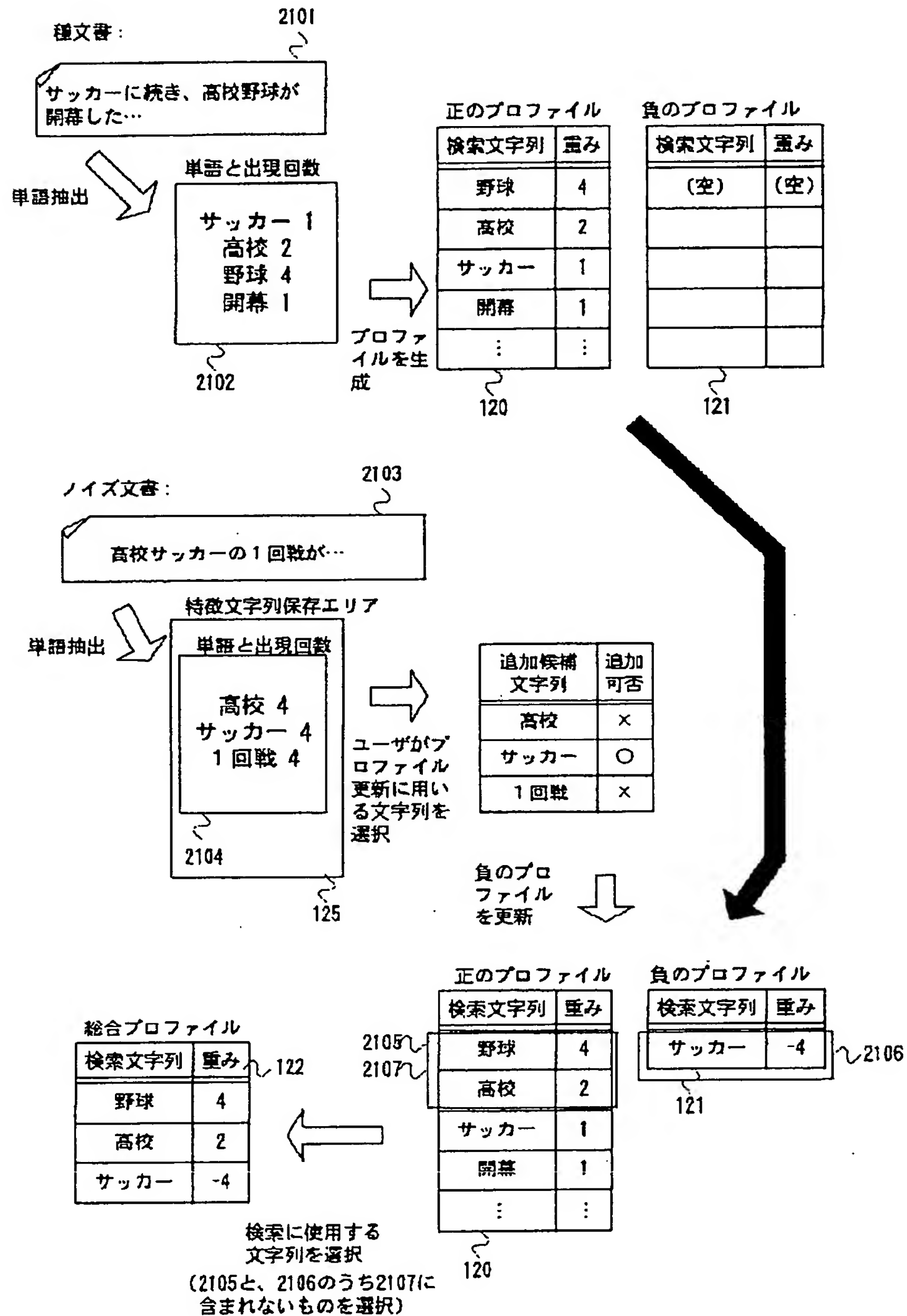
【図 18】

図 18



【図21】

図21



フロントページの続き

(72)発明者 菅谷 奈津子
神奈川県川崎市幸区鹿島田890番地 株式
会社日立製作所システム開発本部内

(72)発明者 松林 忠孝
神奈川県川崎市幸区鹿島田890番地 株式
会社日立製作所システム開発本部内

(72) 発明者 山口 明彦
神奈川県川崎市幸区鹿島田890番地 株式
会社日立製作所システム開発本部内

(72) 発明者 川下 靖司
神奈川県横浜市戸塚区戸塚町5030番地 株
式会社日立製作所ソフトウェア事業部内
F ターム (参考) 5B075 ND03 NK02 NK32 PP30 PQ02
PQ40 PQ46 PR04 PR06 QM08
QS01 QS20 UU06